Proposal for Addressing Grade Inflation at Dartmouth College

11 May 2015

Submitted by the Ad Hoc Committee on Grading Practices and Grade Inflation:

Mark A. McPeek, Biological Sciences (Chair)
Susan Ackerman, Religion; Women’s and Gender Studies; Jewish Studies
Lisa Baldez, Government; Latin American, Latino, & Caribbean Studies
Thomas H. Cormen, Computer Science
Barrett N. Rogers, Physics & Astronomy
Victoria Somoff, Russian

Executive summary

Grade inflation is real. Grades have been increasing linearly since at least the mid-1960’s, and this increase is fueled by various influences. Moreover, the rate of increase is nearly identical in all divisions and departments. The median grade at Dartmouth in Academic Year (AY) 2013–2014 was A–, and 58.7% of all grades given to Dartmouth undergraduates were A or A–. At the current rate of increase, every grade given in every class to every undergraduate at Dartmouth in 2064 will be A. Grade inflation lowers student scholastic effort and deprives them of the realistic academic assessment they need to best further their learning.

This document contains a brief history of grading at Dartmouth, demonstrating that grades have steadily risen over several decades. We propose a solution to grade inflation that is simply stated:

Instructors should assign grades according to the Dartmouth Scholarship Ratings in the ORC.

That is, the grade of A should be given to students who demonstrate excellent mastery of material, the grade of B should be given to students who demonstrate good mastery of material, the grade of C should be given to students who demonstrate adequate mastery of material, the grade of D should be given to students who demonstrate deficient mastery of material, and the grade of E should be given to students who demonstrate seriously deficient mastery of material. Either Dartmouth faculty do not currently grade to these standards, or our courses are insufficiently challenging.

Along with our simple solution, we propose a set of faculty and administrative practices to change the incentives of both faculty and students so that grades are given according to the Dartmouth Scholarship Ratings and students understand the meaning of these grades. We also respond to points that have been made about potential changes to grading practices.
Grade Inflation And Its Effects

The following chart shows the average of all grades given at Dartmouth each year from 1974 through 2014:

The R² value of 0.9901 indicates that these grades fit the straight line shown almost perfectly. Overall GPA has been rising at an average of 0.0109 per year. In other words, the GPA of Dartmouth students has increased at a rate of approximately 0.1 per decade. Given this relationship, we project that every grade given in every class will be an A in 2064.

Here are the grade distributions for all undergraduate classes given in 1974 and 2014:

The median grade has risen from B in 1974 to A– in 2014. In AY 2013–14, 34.0% of all grades given were A, and 58.7% were A or A–. Given current trends we project that 60%
of all grades given in AY 2014–15 will be A or A–. Moreover, A was the most frequent grade given in 44 of the 46 academic departments and programs in AY 2013–14.

Grades are rising in all divisions of the College and at approximately the same rate. Here are the average grades over four decades, by division:

![Graph showing average grades over decades](image)

Although grades in the Arts and Humanities Division tend to be higher than in the Social Science Division or the Science Division, the rates of change—that is, the inflation rates—are approximately the same in all three divisions, with GPAs rising at a rate of 0.0101 to 0.0116 per year. The same trend is seen in the data for each department separately, even for the departments with the lowest overall GPAs. The difference in the rate of change among departments is also slight. Grades have risen everywhere in the College over at least the past 40 years at approximately the same rate.

Do rising grades indicate grade inflation? That is, are grades rising without student performance increasing at a corresponding rate? Or could the rise in grades be explained by students mastering the material better over time? If the students are indeed performing better over the years, might it be because students are actually more adept? Might it be because our teaching methods and technology have improved?

SAT scores for incoming Dartmouth students have been rising, but do they account for the increase in grades at Dartmouth? Here are SAT scores of incoming Dartmouth students since 1979:
Note the discontinuity for the incoming class of 1993, when admissions policies apparently changed with the objective of increasing student academic performance. If SAT scores account for rising grades, then we should see a sharp increase in GPA during the period 1993–1997, as students with higher SAT scores percolated through our courses. These years show only the steady change that we see in other years. No, improved SAT scores fail to explain why grades at Dartmouth have risen.

Not everyone believes that SAT scores predict student academic performance. Moreover, some have hypothesized that the discontinuous increase in SAT scores seen in 1993 is somehow caused by the recentering of SAT scores done in April 1995. Suppose we discount these SAT data and accept the premise that students have become more capable over time, whether by their preparation before matriculation, by their inherent abilities, or by improvements in how we as a faculty deliver content. Do these conditions justify steadily rising grades?

They do not. If grades rise while students become more capable, then our courses are failing to sufficiently challenge our students. If students are more prepared than in the past, or if our teaching techniques have improved, then we owe it to our students to increase the rigor of our courses. Knowledge and concepts that once were considered advanced are now treated as fundamental. When our top students can handle more challenging material, we should challenge them with more rigorous courses.
Indeed, at least one of the following must be true:

We are giving our students higher grades than many of them deserve.

– OR –

Our courses are so non-rigorous that the majority of students can achieve excellent mastery with little effort.

In fact, grade inflation correlates inversely with student effort. For example, a study by Philip Babcock\(^1\) analyzed student evaluations at the University of California, San Diego, and found that students taking classes in which the expected grade was A self-reported spending 50% less time studying outside of class than students taking classes in which the expected grade was C. Data from almost a quarter of a million Dartmouth course evaluations gathered since 2006 agree with this result. Here is a chart showing the percentage of Dartmouth students who expect the grades A, B, C, D, and NR, categorized by the number of self-reported hours per week spent on classwork:

![Graph showing expected grades vs. self-reported hours]

Students who worked 1–5 hours per week expected an NR (non-recording) more than any other grade, followed by A. Students who worked a little more, 5–10 hours per week, expected a B more than any other grade. The most interesting data appear in the rightmost group, where we see that students who worked more than 15 hours per week expected a D more than any other grade, followed by C. Our own data show conclusively that when students expect high grades, they are not motivated to work as hard as when they do not expect high grades.

Based on the above data, our committee has concluded that the faculty can pinpoint the cause of grade inflation by looking in the mirror:

It’s not the grading system.

It’s the graders.

---

Proposal

Guiding principles

To solve the problems caused by grade inflation, we have held to the following principles:

1. Accurate grading, according to the Dartmouth Scholarship Ratings, will increase academic rigor.
2. Every student who demonstrates excellent mastery in a class should receive an A.
3. The course instructor is the only person who can say what grade a student should receive.
4. Differences in the distribution of grades given by different academic units and given in different classes within each academic unit are natural and expected.
5. A mandated distribution of grades enforced by the institution would be antithetical to the academic and educational missions of this or any institution.

The solution


Scholarship Ratings

Regularly Graded Courses: Since the fall term of 1973-1974, the grade assigned at the completion of a course has been one of the following: A, A-, B+, B, B-, C+, C, C-, D, or E. The following guidelines offer general criteria for evaluation and grading, with ‘plus’ or ‘minus’ designations indicating that, in the opinion of the instructor, the student has performed at a level slightly higher or lower than the norm for that category.

A
1. Excellent mastery of course material
2. Student performance indicates a very high degree of originality, creativity, or both
3. Excellent performance in analysis, synthesis, and critical expression, oral or written
4. Student works independently with unusual effectiveness

B
1. Good mastery of course material
2. Student performance demonstrates a high degree of originality, creativity, or both
3. Good performance in analysis, synthesis, and critical expression, oral or written
4. Student works well independently
C
1. Acceptable mastery of course material
2. Student demonstrates some degree of originality, creativity, or both
3. Acceptable performance in analysis, synthesis, and critical expression, oral or written
4. 4. Student works independently at an acceptable level

D
1. Deficient in mastery of course material
2. Originality, creativity, or both apparently absent from performance
3. Deficient performance in analysis, synthesis, and critical expression, oral or written
4. Ability to work independently deficient

E
1. Serious deficiency in mastery of course material
2. Originality, creativity, or both clearly lacking
3. Seriously deficient performance in analysis, synthesis, and critical expression, oral or written
4. Cannot work independently

The following grade point values are assigned: A, 4; A-, 3 2/3; B+, 3 1/3; B, 3; B-, 2 2/3; C+, 2 1/3; C, 2; C-, 1 2/3; D, 1; and E, 0.

In view of the many grades assignable and differences in faculty policies, every faculty member will explicitly declare criteria for grading to students in his or her courses and provide as much information as possible with respect to an individual student’s progress and the evaluation of the final grade assigned.

To solve the problems causing grade inflation, all we need to do is

Offer challenging courses and grade them according to the Dartmouth Scholarship Ratings.

How will we know when our courses are sufficiently challenging and are graded according to the Dartmouth Scholarship Ratings? When the frequency of A declines, and the frequencies of B and C return to intelligible levels, especially in lower-level courses. (Presumably, upper-level courses are taken by students in the context of their major course of study, and students have chosen their major course of study based on their aptitudes and skills and desired career paths. Regardless, even many of our upper-level courses should be much more rigorous than they currently are.)

Some faculty members have suggested that instead of changing our grading practices to conform to the Dartmouth Scholarship Ratings, we should do the opposite: change the Dartmouth Scholarship Ratings to conform to our grading practices. Doing so would only codify the causes and consequences of grade inflation, not solve them. Furthermore, it would not prevent further grade inflation (until we hit the 4.0 ceiling).
It is important to understand what this committee is not recommending:

- We are not recommending any sort of limitation on the number or percentage of any particular grade.
- We are not recommending that instructors or departments set target medians, although some may choose to do so.
- We are not recommending grading according to any “curve.”

What we are recommending is that we offer challenging and rigorous courses throughout our curriculum and grade them so that all high-performing students get high grades, all intermediate-performing students get intermediate grades, and all low-performing students get low grades.

_The best-performing students in a class deserve grades that distinguish them, and low-performing students deserve an honest assessment of their performance._

We propose to accomplish this goal by changing the incentive structure for faculty giving grades. Currently, no incentives exist to foster a faculty member giving low grades to low-performing students, other than the faculty member’s own conscience. On the other hand, many incentives are imagined by faculty to give ever higher grades. Fundamentally, the incentive structure for faculty must be changed to favor faculty giving low grades to low-performing students, intermediate grades to intermediate-performing students, and high grades to only high-performing students according to the Dartmouth Scholarship Ratings.

In order to accomplish this objective, we propose the following specific changes:

---

**We should make public the data we have gathered about grades, past and present.**

If grades at Dartmouth shift downward, then all stakeholders need to know that we have done it and why. Students, faculty, administrators, and the general public—including employers and graduate schools—must understand what grades at Dartmouth will mean. Currently, many students equate a grade of B as “failing” the course. All stakeholders must clearly understand the context of each grade, and that a B is literally a good grade, for it indicates good mastery of the material.

Moreover, in the current Dartmouth environment, where A is the most frequently given grade, receiving an A confers no special signature of achievement on the academic accomplishments of a student in a particular class. If only students who demonstrate excellent mastery get As, then the grade of A will be truly significant.

Conversely, we rarely give Cs anymore. Certainly, we have many students whose work is “acceptable” but not “good.” These students should receive Cs, and it needs to be understood by all that a C is not a failing grade, but instead a passing grade that indicates acceptable mastery.
The NRO should be abolished.

The NRO was introduced in 1967, quite possibly as a way to encourage students to explore courses outside their comfort zone. This reason remains the high-minded justification for this option, but the data belie that students avail themselves of the NR option for its intended purpose. Rather, as the chart on page 5 demonstrates, it mainly provides students a way to drastically reduce their effort in a course. Many students set a high NR target grade and put in very little work, knowing that they can easily achieve the minimum grade of C that does not trigger any academic action yet goes on the transcript as NR; and if they are fortunate, they will receive their target grade. Students will make a greater effort in courses when their grade is their grade.

Course syllabi should include clear statements about the standards used to evaluate student performance and how these standards comport with the Dartmouth Scholarship Ratings.

Students should have a clear understanding of the standards by which the instructor will evaluate their performance in the class according to each category of the Dartmouth Scholarship Ratings. Students would then understand what they need to do to attain each grade. Clear statements about the standards in syllabi will also greatly reduce the amount of questioning about grades at the end of the course.

Resource allocations should be based solely on intellectual and educational merits.

Departments are concerned about the impact of enrollments on resources, most notably faculty lines. High enrollments may form part, but by no means all, of an argument for more faculty lines, but low enrollments should not, on their own, be cause for resources to be rescinded or withheld.

From our analyses, long-term student enrollment patterns—both number of majors and number of students enrolled in courses—at Dartmouth follow national trends. For example, on the next page we present three representative charts showing national trends\(^2\) (red symbols) and Dartmouth figures for the number of majors (green symbols) in Computer Science, English, and History. The number of majors in each of these departments very closely tracks the national trend for that discipline in the number of Bachelor’s degrees awarded at all 4-year institutions in the United States.

Almost every department and program at Dartmouth exhibits enrollment patterns that track national trends (Economics is the only notable exception). The inescapable conclusion from these correlations is that enrollments are driven by national economic and societal interests. Therefore, departments and programs can do little that is academically justifiable to increase their enrollments or prevent falling enrollments.

\(^2\) Data on the number of Bachelors degrees awarded at 4-year institutions in the United States from 1966–2013 from the National Science Foundation Population of Institutions database https://ncsesdata.nsf.gov/webcaspar/TableBuilder accessed 20 March 2015.
Nevertheless, losing resources because of low or falling enrollments was a clear and consistent anxiety we heard from faculty in all divisions that caused many to try to boost enrollments using dubious academic means. We heard stories from many faculty in all divisions about how the fear of low or falling enrollments motivates individuals and entire departments to decrease the rigor of their courses. If the faculty were certain that the Dartmouth administration would not revoke or withhold resources from departments with low or falling enrollments, then instructors would stop reducing the academic rigor of their courses and actually increase the rigor. Grades would then take care of themselves.

We recognize the critical need for more teaching resources in departments with increasing enrollments, either in the short or long term. The Administration should aid these departments, but only if the departments put forward plans to use those resources for a rigorous and thoughtful curriculum.

Additionally, the scholarly vitality of Dartmouth depends on a diverse and expansive portfolio of academic disciplines, and not simply on fulfilling enrollment needs. If we are to continue to build a robust research and scholarly faculty in all areas of the institution, investment will be required in some scholarly areas without regard to undergraduate enrollments.

**A strict minimum number of students required for an undergraduate course should be abolished.**

The consequences of low enrollments in individual courses motivate faculty to reduce the rigor of their courses. Currently, if a course has fewer than five students registered, it can be canceled by the administration, and the faculty member must then teach an extra course in a later term. Our recollection is that this policy was put into place around the post-9/11 fiscal downturn. While it may save a few dollars, it has the deleterious effect of shifting a course to a later year and potentially taking away a Residence term from the faculty member that should be devoted to research and scholarship.

The reasons for low enrollments in courses vary across departments, divisions and even individual courses. Upper-level courses in departments with a low number of majors
may have chronically low enrollments, but these courses still must be taught for those students. Certain language programs may see few students who continue beyond the Language 1-2-3 sequence. Yet, if we have made a commitment to teaching these languages, we must also commit to offering students a curriculum that extends at least through the third year. The stochastic nature of student enrollment patterns also means that a course with moderate enrollments (e.g., 15–20 students on average) in most years will have a rare year where enrollment dips to an unusually low level. Faculty may also misjudge student interest in a topic on the first offering of a new course.

Because these diverse reasons for low-enrollment courses differ across divisions, we recommend that each divisional council develop a scholarship-centered set of policies that maintains teaching effort without incentivizing faculty to reduce academic rigor. Associate deans can then ensure equitability of policies across divisions. We imagine a set of policies that deal with the typical reasons for low enrollments. For example, in departments with lower numbers of majors, low enrollment in some upper-level courses should be expected and so must be permitted, and the associate dean and department chairs should work closely to ensure that teaching effort is equitably allocated among faculty in those departments. A sporadic under-enrolled offering of a typically moderately enrolled class also should simply be permitted. In a major with moderate to high numbers of students, low enrollment in a course on its first offering or for two consecutive offerings may signify that this course is unnecessary and should be discontinued. Faculty whose courses routinely have low enrollments also should be asked to adjust their curricula. Regardless of the specific policy details that are developed, our strong recommendation is to remove the penalty of having to make up courses because of low enrollments, which is the incentive to decrease rigor in order to increase enrollments.

This change is critical, because we heard a diverse set of reasons for low enrollment courses across divisions, but consistent stories that faculty frequently give higher grades than are warranted in order to attract students and thereby ensure that their courses are not canceled. Moreover, we have anecdotal evidence that some faculty encourage students to register for their courses and stay in the course just long enough to avoid having the course canceled. Such gaming of the process is only detrimental to the institution’s educational mission. Removing this incentive to attract students to a course simply to prevent it from being cancelled is the only way to stop such behavior.

Each academic unit should receive a report from the Registrar each year on the grades given within the unit and historical trends.

This report will characterize the frequency of all grades given in each course and an overall department frequency distribution, and the full report will be distributed to every member of the unit. Faculty should discuss the distributions, the standards they set for students to meet the Scholarship Ratings, how the grade distributions reflect the rigor of courses, their experiences with grading, their expectations and thought processes in assigning grades, and goals for the unit to accurately represent student performance to the students.

This discussion should be used to set priorities, articulate standards and strategies for student assessment, and align as much as possible the expectations of the various faculty
according to how they interpret their standards in each category of the Dartmouth Scholarship Ratings. These data will also be used for articulating the unit’s strategies for competing for resources through its academic and educational mission.

The chair of each academic unit should report to her or his Associate Dean at the beginning of the fall term on her or his assessment of course rigor within the unit, time spent by faculty on courses, evaluation methods, grading standards, and grade distributions of the previous year.

These reports are a self-assessment of the philosophy, structure and rigor of the teaching methods and teaching effort within the unit. The grade distributions to be used in these reports are those supplied to the units by the Registrar, as described in the previous recommendation. These reports will then be distributed to the various chairs in the department’s divisional council and discussed at a fall divisional council meeting.

Tenure and promotion dossiers should contain clear statements of the standards used to evaluate students and the grade distributions for all courses taught by the candidate while in rank.

Teaching statements should contain a clear statement of the standards used by the faculty member for evaluating student performance and abilities in classes in each of the various categories of the Dartmouth Scholarship Ratings. This information should be expected for promotion decisions at all levels, and it should articulate the expectations that the faculty member has for student performance for each category of the Dartmouth Scholarship Ratings, as well as how the faculty member evaluates these expectations in classes. The distribution of grades given by the faculty member in each course taught during the period in rank should be included in the dossier that is evaluated by the faculty member’s senior colleagues in their department and by the Committee Advisory to the President (CAP) and used as a metric of the academic rigor of the courses.

The senior members in the department should provide an assessment of the candidate’s standards for evaluating students, and how well the candidate’s standards match unit expectations about student performance relative to the various categories of the Dartmouth Scholarship Ratings as part of their evaluation.

In addition, when considering student evaluations at both the departmental and CAP levels, focus should be placed exclusively on student comments about the effectiveness of teaching, the rigor and academic challenges presented to students by the coursework, the workloads expected of students, and the standards to which students were held. A range of student opinions may also be expected, since lower-performing and higher-performing students may have different opinions about a rigorous course. Whether the students liked the course or the instructor is orthogonal to the education the student received in a course, and so how much the students liked the course or the professor should be immaterial to the deliberations of the department and CAP.

Evaluations of non-tenure-track faculty should follow exactly the same guidelines. Enrollment statistics should not be considered in any way in evaluating non-tenure-track faculty. The evaluation of the faculty member’s teaching should focus exclusively on
course content and rigor, student effort, assessment methods and grading standards, and how well grade distributions reflect the rigor of the course and comport with these grading standards.

At the yearly meeting with their Associate Dean, junior faculty should explicitly discuss how they set standards for each category of the Dartmouth Scholarship Ratings and the grade distributions they have given under those standards.

Junior faculty should discuss the distribution of grades they have given in courses since the previous meeting and discuss with the Associate Dean the standards they use in determining what constitutes various levels of performance in these classes commensurate with the various categories of the Dartmouth Scholarship Ratings.

Each year in the Faculty Record Supplement, faculty members should report their grading standards and the distribution of grades they gave in each class in the three previous years.

Faculty should articulate how they assess students in the various classes they teach and how those assessments are objectively converted into grades based on each of the categories in the Dartmouth Scholarship Ratings. They should explain how their grade distributions identify the academic rigor of the course and comport with these grading standards. These statements and grade distributions should be used as a factor in determining merit raises, based on how their grade distributions align with the stated goals and standards of the academic unit in which the faculty member resides.
Responses To Concerns Raised

The facts and recommendations in this report have been disseminated to some faculty, students, and even the general public. Here we address some of the concerns that have been raised.

“Students won’t get into medical, professional, or graduate schools if we give low grades to low-performing students.”

Lee Witters, faculty advisor for the Nathan Smith Society and faculty director of Pre-Health Advising, kindly obtained data from the Association of American Medical Colleges for our use on all Dartmouth graduates who applied to medical schools for admission between 2002 and 2014. The two charts in the first row below present all the data for 2002 and 2014 as representations of the completeness of the data set. Each symbol represents one student, and its position identifies the GPA of the student when she or he graduated from Dartmouth and the number of medical schools to which the student was admitted that year. In each of the two years, a student with a GPA of 2.4 was admitted to not one, but two, medical schools. Also, in each of the two years, a student with a GPA of 3.7 or above was admitted to no medical schools.

The two charts in the second row plot the maximum, average, and minimum GPA of students who were admitted to at least one medical school (left panel) and who were admitted to no medical schools (right panel) for each year. The maximum and minimum GPAs in the two charts are very close. The average GPAs for those admitted and not admitted differ by about 0.3 each year.
Certainly, having a higher GPA is one contributing factor for successful admissions, but it is far from the determining factor. Admissions committees weigh several components in making decisions: the person’s history of engagement with the medical system, a candidate’s motivation for wanting to enter the medical profession, letters of reference from current medical practitioners and faculty, MCAT results, as well as GPA.

We must also question whether giving high grades to low-performing students in order for them to be competitive for admission to graduate or professional schools is rational or ethical. We believe that the Dartmouth faculty’s responsibility is to offer the best education to all students in order to prepare them for whatever they choose to do in life, and the best way to accomplish this is to offer academically rigorous courses that are graded fairly according to the Dartmouth Scholarship Ratings.

“Giving high grades to poorly performing students will increase my evaluations by these students, which will increase the likelihood that I will be promoted, given tenure, and given higher merit raises.”

We have no data on this assertion, but with the administrative recommendations made above, it will happen only if we let it.

“I don’t want to deal with students complaining about their grades.”

– OR –

“I don’t want to jeopardize the relationship I have with my students by giving them lower grades.”

A faculty member who gives inappropriately high grades to avoid fielding complaints or to maintain friendships with students is simply failing to do her or his job. We must interact with our poorly performing students, and do so in ways that make students more engaged with the course material. Additionally, we must give them clear evaluations of their performance through their grades. As we saw earlier, lower expected grades actually motivate students to work significantly harder in their courses.

Furthermore, low grades in lower-level courses send a strong signal to students that perhaps they should consider a different field of study. When we give students in introductory courses higher grades than they deserve, they can be lulled into thinking that they have an aptitude for the area, only to find out in upper-level courses that they were mistaken.

“Students are already stressed. We hear complaints about student mental health all the time. Won’t making students work harder in their courses compound the problem?”

Stress is a part of life. We cannot remove stress from our students’ lives. Indeed, for most of us, we need some stress to motivate us to succeed.
However, the real question here is what is the source of student stress? We contend that many Dartmouth students overcommit in their extracurricular activities. This is understandable, since they had to compile a litany of activities to be admitted to Dartmouth in the first place. Nevertheless, too many Dartmouth students have too many activities outside of their classes. We wonder whether many students view Dartmouth primarily as a place to live and participate in activities—and, when time permits, to work on their courses. A number of students described to us their time commitments to extracurricular activities that are the equivalent of full-time jobs.

Students need to refocus and understand that they are at Dartmouth to get an education. They have access to world-class faculty, and they need to make the most of their coursework. Extracurricular activities should be just that—extra—what they do in their spare time and not their primary focus.