

# **Report to Dean Mary Miller from the Ad Hoc Committee on Grading**

**February 2014**

## **1 Introduction**

The committee has met a number of times since its report to the faculty was issued in April 2013. Four new members were added, including two students, and one member left. Various members of the committee have met with three student focus groups, divided between STEM, Social Science, and Arts and Humanities. Some members of the committee have met with the Dean's Advisory Committee. A student survey was also done. The chair of the committee met with two faculty departments to discuss grading practices.

The decisions that were made are the following:

We dropped the proposal for numeric grades. Although some members of the committee still like this idea, it seemed too controversial to pursue further. We also discussed the idea of adding an A+ grade in order to distinguish the very top students, but this was rejected as being an awkward way of trying to lessen grade compression at the top. So we are not proposing a change of currency, which means that any lessening of compression must come through the normal letter grades.

We discussed the idea of changing the procedure for choosing Latin Honors. Last year the cutoff for the top 5 percent was a GPA of 3.96. For 5 to 15 percent it was 3.88 percent. And for 15 to 30 percent it was 3.80 percent. Although these cutoffs are high, where a few B's may eliminate a student from Latin Honors, the majority of the committee felt that the objectivity of this procedure dominated other procedures that would require at least some subjective judgment.

We discussed whether more information should be put on a student's transcript. This meeting was well summarized by Josh Kalla in an email, which is as follows. "The registrar, Gabe Olszewski, presented the Dartmouth/Indiana/McGill model of contextualized transcripts and then gave rather convincing reasons why it is not a good idea. First he noted that these schools are transitioning back to the standard transcript, so it seems to be a failed experiment. More substantively, he thinks that simple is better because based on his experience, many people already misread the simple transcript. Adding more information will only increase the confusion.

He also pointed out that it doesn't necessarily provide valuable information. For one, how do you compare a GPA 3.8 student (A) with a contextualized transcript to a GPA 3.8 student (B) with a simple transcript? You might learn that A comes from a school with grade inflation, but B could come from a comparable school, a school with worse grade inflation, or a school with grade deflation. When it comes to comparisons across schools, a contextualized transcript doesn't actually help, unless every school made the switch. He also said that the vast majority of Yale transcripts are viewed by a small number of people - grad admissions committees, fellowship committees, and large firms (Teach for America, finance and consulting). These folks already have an understanding of what a Yale A means based on looking at lots of Yale transcripts over the years. They essentially already have the context. Everyone present at the meeting was convinced by Gabe that we should not pursue this."

We also discussed the grading of independent study and are proposing a change. This change is discussed in Section 4. It will require a faculty vote.

The remainder of the discussion of the committee was on what was broadly defined as "grading theory" and on guidelines. Section 2 has been drafted by Paul North, which summarizes the committee's views on grading theory. Section 3 has been drafted by R. Shankar. It summarizes the committee's views on guidelines.

One of the main points that came out of the student discussions, which is discussed in the next section, is that students would like more grading feedback. Simply giving letters, even if the letter is an A or A-, with no further feedback is not satisfying. It does not enhance the learning experience.

On guidelines, the committee last year wanted all four of its proposals voted on together. It was concerned that if only the first two proposals were passed, in particular grading distributions of departments being known by other departments, with no guidelines, this might lead to harsher grading departments becoming easier in order to compete for students. The data collected by the committee were a shock to many people, in particular data showing that there were departments in which very few B+ grades and below were issued.

In the year since the grade distributions were made available, there has been an increase in the percent of A's and A-'s issued in Yale College. In the fall of 2013 the percent was 68.3 versus 67.2 in the fall of 2012. In the spring of 2013 the percent was 69.5 versus 68.7 in the spring of 2012. These changes are large compared to previous yearly changes. These changes are consistent with the committee's concern that more information with no guidelines might lead to higher grades, but these are, of course, only two observations. (1.5 should probably be subtracted from each of these percents to adjust for credit/D/Fail having fewer A's and A-'s

than the visible grades.)

As Section 3 discusses, the committee still feels that guidelines are necessary if any meaningful change is going to come about. Departments cannot make changes by themselves; they need a frame of reference. The committee also feels that the guidelines proposed in Section 3 are more palatable than those proposed in the original report. The committee feels that Shankar's discussion puts the guidelines in a nice perspective. Section 3 is written in the first person. Shankar presented his report to the committee in this way. The first person has been retained because the perspective would otherwise be somewhat lost.

Section 5 is an addendum by Scott Stern.

## 2 Reflections on the Purpose of Grades

Over the nearly two years of its work, the committee has broadened its thinking beyond the problems of grade inflation and grade compression in Yale College. While we still feel strongly that targeted measures are needed to halt the upward drift of grades across the college, in the course of numerous conversations, public, private, and among the committee, we have come to believe that a shift in the ways teachers, students, and administrators think about grades will be extremely beneficial to the community as well. A thoughtful reevaluation of what grading is for, we believe, will help put grades in their place in the process of higher education and as a result grading may become better integrated into the projects of teaching and learning. We also think that clarifying and also to a certain extent moderating the role of grades in college life will have its own effect on the problems of inflation and compression.

One sign that you have received a grade, as opposed to some other token, is that you don't fully understand what the single letter represents. Does it evaluate the work according to some standard? Does it compare one student's performance to the rest of the student group? Does it express a relative position on the teacher's own scale, on the scale of a particular discipline, in the student's own development? Or is it to be taken in an absolute sense? We believe a grade always has multiple possible audiences and frames of reference for understanding it. On one hand, this is why grades are needed. They enable a translation of a unique event across disparate fields with incompatible standards of value. A grade makes comparison possible at the most general level, but it does so at a cost. All of the detail, all of the differences within and between students, teachers, assignments, types of courses, and modes of knowledge are collapsed into a few abstract letters.

Here is the original list of the purposes of grades from the committee's 2013 report:

“Grades serve educational purposes by signaling to a student strengths and weaknesses in performance. Grades on individual assignments give guidance as to what needs more attention, what can be used to build on, what needs change, and what bears repetition. They point to talents and capacities. Course grades signal possible directions for further study, for selecting majors, and for pursuing graduate or professional study. At the same time that grades give guidance, they also have a motivational function by encouraging, or not, activities of various sorts. They may elicit greater effort or discourage it.

Grades also serve the selecting and sorting functions of society by signaling to outside entities the strengths and weaknesses and overall capacity of individual students. They serve as certifications of competences, both particular and general.”

We still believe that condensed, easily comparable information gives a quick reference for students and institutional actors for these purposes: motivation, improvement, and ranking. And yet, the intense condensation and the inexact comparisons that grades imply are also counter-productive in many ways. This gives rise to confusion among all parties as to whether grades are absolute or relative, and if relative, to which framework or standards.

The ambiguity of grades, we believe, cannot be completely abrogated. It is in the nature of this sort of abstract communication with multiple audiences, frames of reference, and standards that in order to achieve its comparative purpose, its ambiguity has to increase. This is why we would like to emphasize that grades are first and foremost communicative acts. Giving a grade is an act of communication between several different pairs of interested parties at once: teacher and student, teacher and advisors, a student and her parents, a student and external institutions, the institution and itself. Of these, the teacher-student communication is primary. But this is also the exchange in which grades mean the least as compared to other forms of exchange. Indeed, unless they are encased in expansive, analytical commentary, grades remain a shorthand and may even misrepresent the complexity of the learning event.

Because grades, being abstract and general, are partly at cross purposes with their pedagogical function, the committee feels that grades should be treated as a secondary form of communication. In breadth and specificity, grades fall well short of the primary ways teachers and students interact: the syllabus, assignments, written commentaries on student work, lecture, class discussion, office hour chats. The committee feels that grades should be meaningful, but they can only be so in the context of frequent, richer discursive communications. If they are to be meaningful to teacher and student, grades should faithfully represent a conscientious evaluation of the coursework, according to explicit standards. Since the student is not privy to the judgments in grading, what the grade actually represents needs to be fully explained.

A grade given by a teacher to a student should be embedded in the fullest possible context. Teachers should communicate in the syllabus the weights of different assignments for the final course grade. For individual assignments, the context, we believe, consists of two aspects: rubrics and frames of reference. A

rubric tells the student what elements of an assignment are being evaluated and what poor, fair, good, and excellent work on those elements would look like. We hope teachers will also communicate the multiple frames of reference against which a grade makes sense. For example, an A± may be relative to a student's earlier work in the course, or it may be relative to the achievements of the other students in the same semester or over the course of several years, or it may be relative to a body of knowledge that has to be attained in order to enter a discipline, and it is often a mixture of these. A grade is an effective communication when the rubric for individual assignments and the frames of reference are communicated with it. We heard from many teachers who communicate both these things before students begin to work on an assignment, and some communicate these things in general terms at the beginning of a course or in a syllabus.

In sum, grades are a secondary communication that without thorough explanation are often confusing and cause anxiety. For this reason, and taking into account the important feedback we've received from students and faculty, we make the following recommendations:

For teachers:

1. Please make your grading policy, the standard and the frames of reference, clear for each course you teach and for each assignment where appropriate.
2. Encase grades, wherever possible, in richer, more thorough prose explanations of the achievements and possibilities for improvement.
3. Give more frequent grading instruments, so that students get an ongoing experience of the grading policy in action, and also so that they have a chance to improve from the comments.
4. Return assignments, even final assignments, with commentary.
5. Some form of written commentary at the end of a course is appropriate. This should give the teacher's overall view of a student's progress and possibilities in the subject.
6. State explicitly what "participation" consists in, if this is part of the final grade.
7. Make use of the "End of Term Report" in the faculty grading system to signal exceptional qualities to the student, dean, and future institutions.

8. One of the most frequent student complaints is about inconsistency in grading across discussion sections. Train TAs in the course's grading policy and institute any other procedures that help insure consistent grading of assignments between TAs.

For students:

1. Please remind fellow students that grades should be treated as secondary to learning.
2. Where a grading policy is absent or unclear, speak to the teacher and request a fuller account.
3. In order to reinforce that natural language commentary is more useful and fairer than an abstract letter grade, read teacher comments carefully and respond to them where you can.
4. Ask teachers for final assignments where they are not returned, and request fuller explanations of grades when you don't understand.

### 3 Grading Guidelines

Our committee was formed two years ago to address the problem of grade inflation, compression, and non-uniformity of distributions across Yale. There is broad consensus in the committee, consistent with Dean Miller's appointment of the committee, that the problem is real and urgent.

My earlier view was that a quantitative problem like this called for a purely quantitative solution. However my views have evolved, largely by listening to fellow committee members. I am always happy with the final outcome when that happens in a committee.

More specifically, I started strongly in favor of the following concepts: (1) A numerical grading system 0-100. (2) Guidelines that specified the percentage range of every grade (A through D) in every course

I had not given much thought to how exactly I graded and had not discussed grading practices with my colleagues. I had a philosophy but it was not spelled out in detail even to myself, let alone my students. I did emphasize, however, that I would not grade on a curve and that if the entire class deserved an A or an F, it would get it. The main point behind this was to reassure the students that the success of their peers did not in any way endanger their own, that cooperation had no adverse consequences if engaged in responsibly.

Here is my current view, as we come to the final stretch. (1) Instructors should be encouraged to think more about how and why we award grades and share them with the students. (2) While this will have many benefits, and could even ameliorate the inflation and compression, we must propose quantitative guidelines that are constantly in the air to keep the issue alive in the minds of instructors and bring about the solution towards which the whole community gravitates. A one shot invitation at soul searching will not provide a continued and complete remedy to the problem at hand.

Here are the proposed guidelines and how they are to be understood:

1. The percentage of A's plus A-Minus's be 40/50/60 for introductory/intermediate/ advanced classes.
2. The instructor can decide the level of the course in ambiguous cases.
3. While guidelines are merely suggested and not mandatory, some numerical ball-park figure is essential if the university as a whole is to strive for some common measures of excellence.

4. The guidelines are to be applied ON AVERAGE: averaged over time, over many batches of the same course and over different courses.

It is left to the discretion of the instructor to account for real and inevitable variability of class quality. These numbers (40/50/60) are not derived mathematically and cannot be defended as such. They are the averages I found in my own records over the years. I found that on some years, I would be substantially off these averages. The only certain thing that can be said of these numbers is that they are in the ball park that would rein in the current run-away situation without being draconian. It should also be noted that in smaller upper level courses, changing the number of A/A-minuses by a small amount can change the percentage a lot. Thus there is an increase in the intrinsic statistical uncertainty of these targets as we move up in the level. I do not imagine the instructor staring at the guidelines while grading, or forcing a curve on the class, but rather, consulting the guidelines after grading to ensure that serious deviations are attributable to reasonable causes.

There was some discussion in the committee about whether the guidelines should be ranges, like 35-40/50-55/60-65. This was rejected because it is actually more restrictive than simply using one number per type of class. Simply using one number per type of class allows people to deviate as much as they want from the proposed average or mean numbers 40/50/60 if they feel it is justified on a case by case basis. To give a window like 35-45 (say for the elementary courses) would specify the mean and the deviations, and would make sense if there were some penalty for going outside the range, in a system where the guidelines were mandatory. Right now what it would do is suggest that the deviations from the mean are expected to lie in the stated intervals. Instead what we propose are specific values for the mean, leaving totally free the deviations ON ANY SPECIFIC CASE. Of course if a professor finds his/her grades are above or below the mean by 15 percent five years in a row, he/she needs to recalibrate.

## **4 Proposal for Grading of Independent Study**

### **PROPOSAL:**

Grading for independent study be changed from letter grades to written reports. The letters “R” or “F” would be recorded on the student’s transcript, and the report would go into the student’s file. (“R” is just an arbitrary letter denoting “report.”) Independent study courses would still count for credit. The overall GPA would not be affected by the report unless the student failed, in which case the F would go into the overall GPA. The letter "R" would not be recorded on the transcript until the report was received from the instructor.

### **REASONS FOR THE CHANGE:**

Independent study courses present special challenges in two ways: (1) the opaqueness of the course title (e.g., Directed Reading, Independent Study, Special Project, Special Tutorial, etc.) AND of the letter grade means that no real information is being conveyed about the nature of the work involved in the course nor about the student’s performance; and (2) such courses do not provide a community of learners as exist in other courses where grades can have a fuller meaning.

When we add to this the fact that grades in these courses have been more than 91percent A or A- in the past two academic years, the members feel strongly that such courses provide a case where detailed feedback should not just contextualize the letter grade but replace it.

## 5 Addendum by Scott Stern

The above report has been thoughtfully considered and carefully crafted. Hopefully, it will evince a positive change for the culture of grading at Yale. Yet I would like to very briefly register some concerns that I believe I share with many students across campus. First, we reject the idea that grade inflation is a "problem" to be solved by harsher methods or restrictive guidelines. Grade inflation, while undoubtedly an existent phenomenon, is the natural result of an improved caliber of students, the increased importance attached to grades, and the altered nature of college itself. Few if any professors would deny that the quality of new admits to Yale has risen over the past few decades; certainly, standardized test scores have gone up. At the same time, rates of admission have plummeted. In response, students are working harder and harder to get into college, and, once they arrive, they must continue to work hard to keep up. Unlike in decades past, "gentleman's Cs" simply won't cut it, and Yale students are aware of this. More and more of them perform A-work because that's what they learned to do in high school, because they need to do so to stay competitive in an unstable economy, and because they must compete in a cutthroat job market.

Even if this weren't the case, inflated grades at Yale would still be a positive for the campus. Inflation allows students to be non-competitive with their peers, even as they operate in a system that, at other colleges, demands a mean-spirited and competitive attitude. Unlike at peer institutions, grades are not a cause for enmity at Yale. They certainly remain a cause for concern - students stress and worry more than professors can know. Yet with Yale's relatively high grades, students do have time to pursue a vast array of extracurricular activities, which enriches our campus and makes our undergraduate experience rivaled by none.

I fear that attempting to alter the grading policy at Yale will mess up this entire, beneficial system. It will stress out and force into competition students who believed they were at a less competitive place. It will dissuade prospective students from applying to a school that they know has intentionally lowered its quality of intellectual life - as Princeton has. It is attempting to fix a problem that isn't truly a problem. Please do not forget the student voice when you do with this report as you will. We are neither entitled nor spoiled, but many of us are terrified that the campus we all love will be changed for the worse. This should not happen.

Nonetheless, I understand that the recommendations of this committee are a reality, student objections aside. In light of this, I think it is vitally important to stress one more thing: the recommendations proposed by this committee are optional. While it may not be entirely apparent upon reading this report for the

first time, these recommendations are just that: recommendations. Professors are entirely free to disregard them if they want. This is essential - faculty will revolt if they feel their academic and curricular freedom is being restricted. If the guidelines are intended to get faculty thinking about the grades they give, that is probably a positive. But if they are meant to be a prescriptive checklist for nervous faculty to adhere to, they are profoundly wrong. Faculty must be made to understand that these guidelines are, fundamentally, optional.

## **6 Committee Members**

Russell Adair, Office of Institutional Research

Ronald Eyerman, Department of Sociology

Ray Fair, Department of Economics, Chair

Robert Harms, Department of History

Roger Howe (fall only), Department of Mathematics

Joshua Kalla, Morse College, Class of 2014

Tina Lu, Department of East Asia Languages and Literatures

Paul North, Department of German

Mark Schenker, Dean of Academic Affairs

R. Shankar, Department of Physics

Scott Stern, Branford College, Class of 2015

Ruth Yeazell, Department of English