How Much Information is Conveyed by a Princeton Grade?

Edward W. Felten
Department of Computer Science
Princeton University
April 22, 2004

Introduction

One of the standard arguments in the grade inflation debate is that grades convey less information than in the past. Because grades tend to cluster at the top of the scale, the argument goes, grades do less to distinguish the exceptional student from the very good, or the very good from the mediocre.

In this note, I evaluate that argument quantitatively, based on historical grading data supplied by the office of the Dean of the College. Using information theory, a branch of theoretical computer science, I compute the average information content of a Princeton grade, and how that content has changed over the last thirty years. I find that a Princeton grade conveys about 11% less information than it did thirty years ago. Adopting a 35% quota on A grades would, at best, increase the information content of grades by 10%.

Grades and Information

Before crunching the numbers, let me take a moment to explain what I mean when I speak of the "information content" of a grade. In short, a grade conveys information by telling us how a student's performance compares to that of other students. It does this by dividing students' performances into groups, where the groups are in ranked order, so that we know, for example, that every student in the A- group performed better than every student in the B+ group (and all lower groups), but we cannot distinguish the performances of students in the same group.

To quantify the information content of a grade distribution, I use a quantity from information theory known as the Shannon entropy, named for Claude Shannon, who received an honorary degree from Princeton in 1962. This quantity is widely used as a measure of information content. Rather than reiterating the formal definition of Shannon entropy, I will instead use a few examples to provide some intuition about how it works.

Suppose we decided in advance that every student would receive an A in every course. Then if you learned that Alice got an A in my course, that fact would tell you nothing at all about Alice's performance. In a system where all grades are equal, grades convey no information at all.

Now suppose that our grading system had two possible grades, A and B, and that we gave an equal number of A's and B's. Then a grade would convey one binary "bit" of

information, because it would allow us to divide students' performances, in binary fashion, into two equal-sized groups.

If the groups were lopsided, so that (say) 75% of students received an A and 25% received a B, then a grade would convey less information on average, because most of the students would be in the relatively undifferentiated A group. According to the formula, a grade would then convey about 0.8 bits of information.

The formula for information content obeys two general rules. First, all else being equal, more information is conveyed if we divide students into a larger number of groups, rather than a smaller number. Second, more information is conveyed if the groups are more nearly equal in size.

An "Optimal" Policy

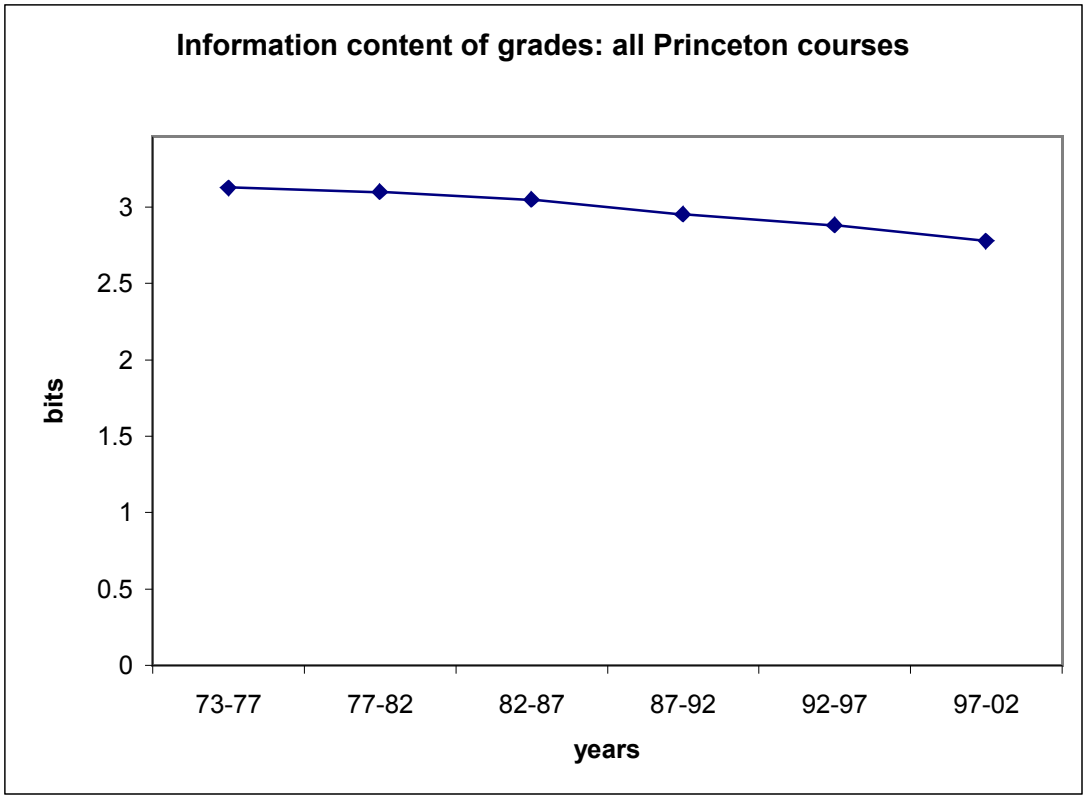
Given our current grading scale, with eleven possible grades ranging from A+ down to F, we could maximize the information content of our grades by giving each of the eleven possible grades to equal numbers of students. Thus one-eleventh of the students would receive an F, one-eleventh would receive a D, and so on up to the lucky one-eleventh who received an A+. I will call this the "fully deflated" policy. Under this policy, each grade would convey 3.46 bits of information.

Of course, nobody would seriously advocate adopting the fully deflated policy. It is useful only as a limiting case, and to remind us that information content is not the only worthwhile criterion for evaluating grading policies.

Analysis of Princeton Grades

To determine the information content of Princeton grades, I used data from Dean Nancy Malkiel's memo to the faculty, "Truth in Grading: Proposals and Questions," dated February 20, 2003. The memo gives grade distributions, for historical periods from 1973-77 through 1997-2002. I computed the information content of grades (i.e., the Shannon entropy of the grade distribution) in each relevant period. I included all course grades in Princeton courses, except grades of P (which convey little if any information in any case).

The graph below shows the results for all Princeton courses.



As the graph shows, the information content of a Princeton grade has declined, from 3.12 bits in 1973-77, to 2.77 bits in the most recent period. This is a relatively modest decline of about 11%.

Perhaps more surprising is the fact that the information content of grades, even today, is not far below the value for the fully deflated policy (3.46 bits), which would be at the top of the vertical scale in the graph above. Our grades today still convey about 80% of the information that they could convey if we adopted the (unrealistic) fully deflated policy.

Effect of a 35% Quota on A's

A proposal currently before the faculty, if adopted, would require that no more than 35% of grades in Princeton courses were A's. About 45% of grades were A's in the 1997-2002 period, so a 35% quota would push about 10% of grades from the A range down into the B range. It is less clear what would happen to those students currently in the B range. Presumably a few would be pushed down into the C range, but we can only speculate.

At worst, the result would be a distribution with many students bunched in the B+ and B grades. This might possibly reduce the information content of grades, though that outcome seems unlikely.

At best, we would return to the grade distribution of 1982-87, when we last gave less than 35% A's. If this is the outcome, then adopting the quota would have the effect of increasing the information content of a grade from its current value of 2.77 bits to the 1982-87 value of 3.04 bits – about a 10% increase in information.

Single Course Grades versus GPAs

A 2002 article in Slate magazine, by Prof. Jordan Ellenberg of Princeton's Department of Mathematics, suggests that even if grades in individual courses do a relatively poor job of discriminating between different levels of student performance, students' grade point averages, which are computed by averaging over a large number of courses, may still enable precise comparisons. (Prof. Ellenberg's article is available at http://slate.msn.com/default.aspx?id=2071759) In other words, a 10% difference in the information content of each individual course grade may lead to a much smaller difference in the information content of a student's GPA.

Conclusion

Princeton grades convey about 11% less information than they did thirty years ago. Adopting a 35% quota on A grades would, at best, increase the information content of grades by 10%.