Letter from the Editor-in-Chief:
Are Grades a Necessary Evil?

Russell L. Herman
The University of North Carolina Wilmington, Wilmington, NC

No matter how much we employ innovative teaching methods, investigate how students think and learn, or spend late hours networking with our classes, eventually most of us will have to grade exams and papers, post grades for courses, and fret over recording less than optimal grades for some students. Grades can disappoint both students and faculty. They have also been the subject of debate within and outside academic circles.

Are grades necessary? There is an expectation that end-of-course grades are useful in providing information about student achievement. Some instructors say that students should come to class with a desire to learn and not have to stress over grades. After all, the university is a community of scholars and introducing grades into the mix is contrary to our goals. Other instructors would attribute some aspect of importance to grades. Grades are used for assessment; as feedback for both students and instructors; for evaluation of student performance; as motivation to learn; to increase competition amongst students; or, to rank and sort a cohort of students. The grades are then be used by parents, employers, or post-graduate institutions to open doors for further education. Setting high standards leads to a trust in the professional we count on such as doctors, nuclear engineers, chemists working with drugs, bridge engineers, and teachers.

Current grading practices are relatively new. The first public exams, such as apothecary or civil service, started in the 1800's. By the 1900s, school systems had grown and the report card became a common instrument for reporting student progress (Moll, 1998). Grades were often reported as percentages. During the 20th century school systems experimented with different grading systems. Now we mostly see letter grades with the possible additions of plus/minus, pass or fail, and possible addition or deletion of some subset of these. Interesting discussions of this history can be found in Drum (1993) and Schneider and Hutt (2013).

Grading is personal and the reasons for assigning and weighting grades are up to the instructor. The composition, or what goes into a grade, could be based entirely, or in part, on examinations, effort, practice, and improvement. The University of Illinois at Urbana-Champaign Center for Teaching Excellence (2009) is one of many centers which provides advice for all aspects of grading. Having decided on the approach to take towards grading which is appropriate for the discipline and level of the course, the instructor

\[1 \text{ Author's email: hermanr@uncw.edu} \]

\[2 \text{ The difference between letter grades and percentages is referred to as norm-referenced or criterion-referenced, respectively (UNC Center for Faculty Excellence, 2012).} \]
should elaborate on the grading rubric in the syllabus. Covington (1989) discusses the meaning of grades. However, we need to understand that grades shouldn’t be “confused with self-worth or innate intelligence.”

Determination of which slot to assign to students is an art. Those who believe that the grades should be distributed such that a few students are deemed excellent and others should fail might use a bell curve. However, there can be problems stemming from the misuse of the bell curve. [See Fendler and Muzaffar (2008) for an interesting discussion of the history of the bell curve.] Allain (2009) and Richeson (2008) describe in their physics and mathematics blogs, respectively, issues and techniques surrounding curving exam grades. Other instructors might allow all students to receive high grades or all to get low grades depending if they have accomplished the learning outcomes spelled out for the course. Administrators sometimes discourage such practices like in the case that a biology professor was removed and her grades were adjusted higher (Jaschik, 2010).

No matter how one has decided to determine end-of-course grades, a final grade distribution will be recorded. It is these final distributions that have been used by others, rightly or wrongly, to glean something about the history of grading practices over time to compare groups of courses, instructors, or universities to each another. As one example, former Duke University professor Stuart Rojstaczer (2002) analyzed data from studies of grade inflation that had been compiled on many schools and had reported these results at his web site, http://www.gradeinflation.com/, in interviews, and later in articles (Rojstaczer & Healy 2010, 2012). As noted by gradeinflation.com,

“In the 1930s, the average GPA at American colleges and universities was about 2.35, a number that corresponds with data compiled by W. Perry in 1943. By the 1950s, the average GPA was about 2.52. GPAs took off in the 1960s with grades at private schools rising faster than public schools, lulled in the 1970s, and began to rise again in the 1980s at a rate of about 0.10 to 0.15 increase in GPA per decade. The grade inflation that began in the 1980s has yet to end.”

Rojstaczer and Healy (2010) had noted that there was a 0.12 increase in GPA in public institutions and 0.15 in private institutions over a fifteen year period ending around 2005-6. These results have become referenced in the media since that time. However, one needs to be careful as to what these results mean. Hu (2006) notes that grade inflation consists of a number of pieces: grade increase, grade inflation, grade compression, and grade disparity. While grade increase is simply an increase in grades, grade inflation refers to an upward shift in GPA without a similar rise in achievement. One must also take into account grade compression, which occurs when grades increase but differences in student performance are hard to distinguish at the high end of the grading scale. Grade disparity results when different professors award drastically different grades the same level of student achievement. Therefore, the sole reporting of grade increases is not enough to indicate that there is grade inflation. [See (Johnson, 2006) for a review of Hu’s report.]

3 Similar discussions are reported in the U.K. as noted in the Daily Mail, http://www.dailymail.co.uk/news/article-2238151/University-exams-enter-history-books-coursework-blamed-grade-inflation.html

©2013 All rights reserved.
As an example, in 2007 we had looked at grade distributions on our campus. In Table 1 we see that there is no noticeable trend in the GPA’s. We then looked at grade distributions for lower and upper level courses and played with the data. As one example, we considered distributions of all grades, A, A-, B+, B, B-, etc. and then regrouped the data into bins of A’s, B’s, C’s, D’s, and F’s & WF’s by combining the B-, B, and B+ grades, etc. In Figure 1 we show these trends for the data we had at the time. We found that the C’s dropped an average of 0.43% per year, while the combined A and B grades rose about 0.42% per year. The combined A+B+C grades are about level, which agrees with the apparent constant percent of both D’s and F’s in the figure. Thus, the change grade distributions are not reflected in the average GPA.

Table 1. Academic Year GPA’s from 2001-2007

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>GPA</th>
<th>Academic Year</th>
<th>GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001-02</td>
<td>2.778</td>
<td>2004-05</td>
<td>2.855</td>
</tr>
<tr>
<td>2002-03</td>
<td>2.826</td>
<td>2005-06</td>
<td>2.847</td>
</tr>
<tr>
<td>2003-04</td>
<td>2.866</td>
<td>2006-07</td>
<td>2.839</td>
</tr>
</tbody>
</table>

Figure 1. Trends in Grade Groups A, B, C, D

However, noting a rise in SAT scores, the slight upward trend of grades, and the short time span, there was no data confirming that the rise in A’s and drop in C’s, or the lack of a constant upward trend in GPA, were strictly due to either grade inflation or the changing characteristics of the student population. Nonetheless, in recent years some universi-
ties have are fighting grade inflation by contextualizing grades on student transcripts; i.e., included the class average for courses (Eubanks 2011).

Grading is one part of what we do. While grades are not acknowledged to be part of program assessment, the way in which one evaluates student achievement can be used effectively to provide a better learning environment. Grades can be used to evaluate the quality of student work, communicate that quality with students and others, motivate students, and guide instructors in the design and organization of courses. While there is not one way to handle grading, an instructor can put as much effort into the grading process as the topics and assignments to be covered. An instructor can set learning goals, assignments, and tests to measure what is valued in each topic and expectation of what the students should be learning. Instructors can then assess their grading practices. [See (Wallooord & Anderson, 1998) for ideas on effective grading.]

Do not leave the motivation up to the pure motivation of students needing to attain a certain grade, but seek to motivate learning and tie achievement to that. Look at assignments and exams and evaluate them as to whether or not they are effectively assessing the achievement of your goals or providing sufficient motivation for the students. Some of the ways to do this can be found in the writing of many of the authors in JET. As noted by a colleague, grades are not a measure of catching students at their worst, but should be used to motivate students and instructors to do their best work.

References


©2013 All rights reserved.


