

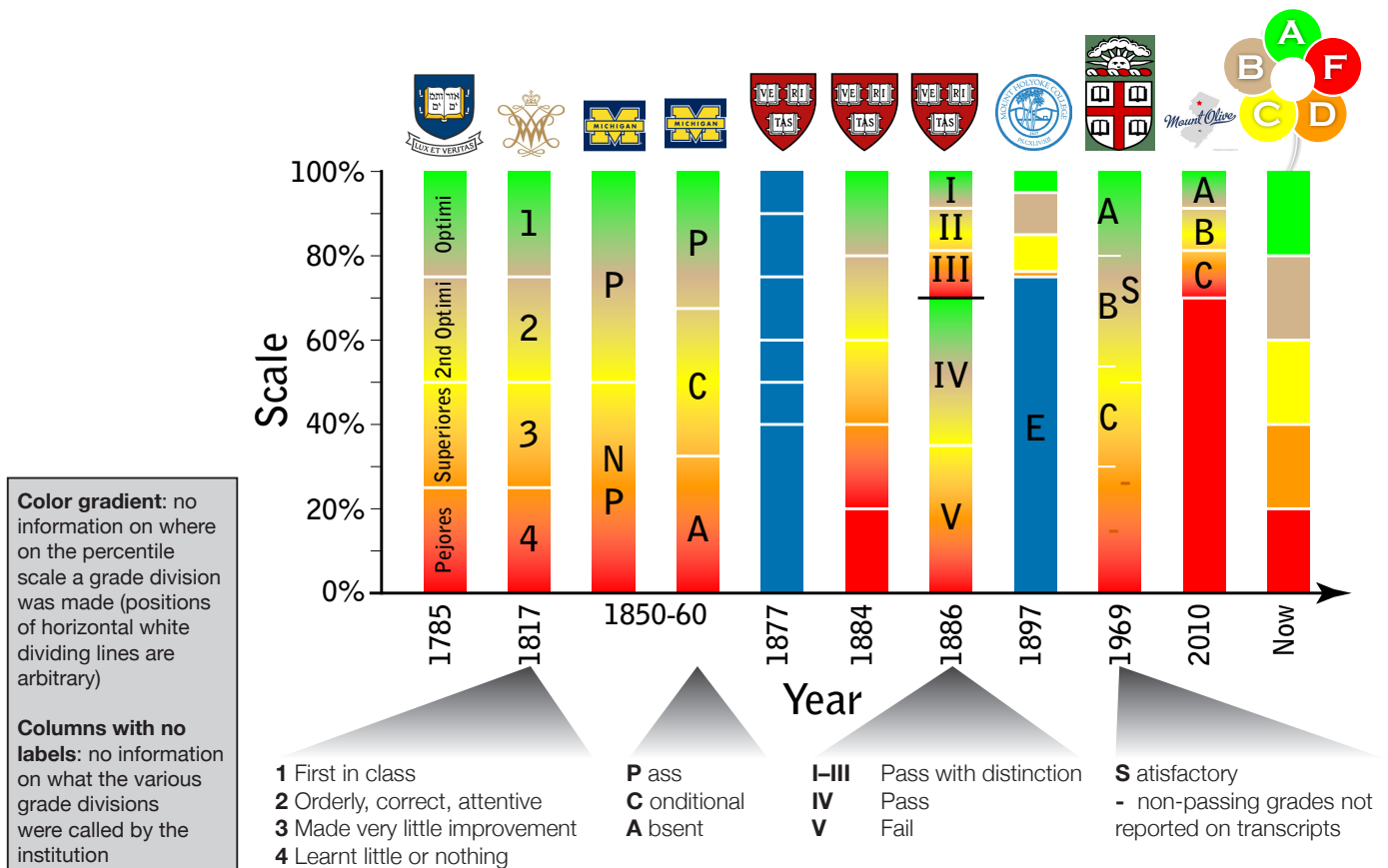
## Discouraging academic dishonesty in online exams using Bloom's Grading

Joseph Ross, Ph.D.  
Associate Professor of Biology  
California State University, Fresno

[jross@csufresno.edu](mailto:jross@csufresno.edu)  
[@rossbiology](https://twitter.com/rossbiology) (Twitter)  
[eduproffer.blogspot.com](http://eduproffer.blogspot.com)

With the advent of more courses moving to virtual delivery, particularly during the COVID-19 era, I've had more frequent conversations with faculty about how to prevent cheating during online exams. Even before COVID-19, I regularly administered open-internet exams in my face-to-face courses. Two simple changes to assessment design and to grading produced more cheat-resistant assessments by assigning more points to exam questions requiring higher-level Bloom's taxonomy (Anderson and Krathwohl 2001) levels, like Analyze, Evaluate, and Create. These are inherently more cheat-resistant activities, because they tend not to have answers or responses that are easily found by search engine query. Student responses to such prompts are also expected to be unique, resisting plagiarism.

Letter grading arises from our tendency to organize and categorize. Somewhere along the way, categorizing learning by ranking students, either against each other or against an instructor standard, became the norm. According to Durm's (1993) description of the history of letter grading, Yale was the first university with evidence of such a system, where students were graded into four tiers. Higher education institutions subsequently tinkered with grading concepts, including the introduction of the pass/no pass concept, a five-tiered system based on a percentile scale, and the introduction of the A/B/C/D/E scale. At some point, E disappeared and was broadly

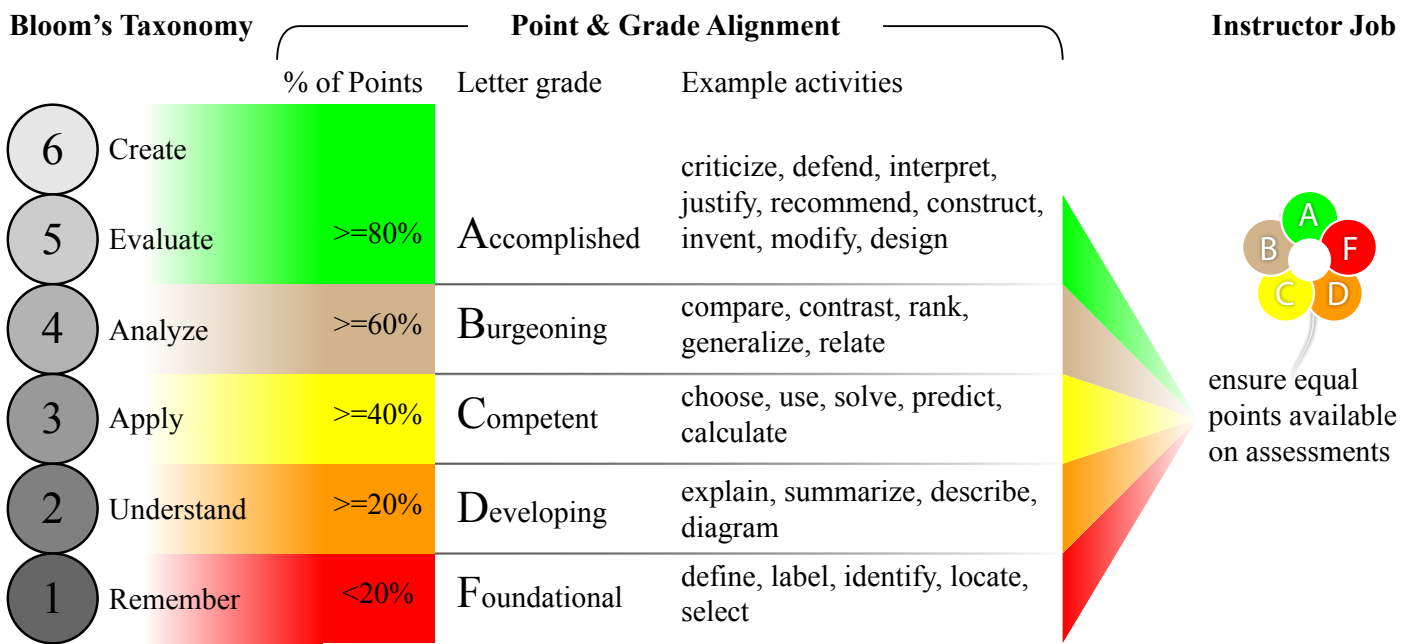


replaced by F. This history of letter grading makes one point clear: we are not inextricably tied to 60/70/80/90% point thresholds for students earning D/C/B/A letter grades.

Like many aspects of grading, the relationship between points and letter grades is entirely arbitrary. Multiple factors interact to produce a single letter grade for each student in each course, usually integrating multiple factors of a student's performance. The myriad variables under the instructor's control include the types and relative weights of activities students must perform and the percentile thresholds separating letter grades. Let's manipulate these to discourage academic dishonesty.

In the USA, it is common (and arbitrary) for 0-59% of points to equate to an F, 60-69% a D, 70-79% a C, 80-89% a B, and 90-100% an A. Why do we compress "passing" letter grades (D through A, or perhaps C through A, depending on your program) into such a small percentile range? When I started designing exams, this traditional practice indicated to me that I should create about 60% of the points on assessments that would be easy for many students to earn. This would ensure that they would get at least to the D level (in many courses, 60% of points) and hopefully then produce a bell curve of letter grades by using the remaining 40% of points to carefully (and still arbitrarily) distinguish D, C, B and A students.

Instead, I evenly distribute letter grades across the entire percentile range, with 20/40/60/80% score thresholds for earning D, C, B and A letter grades. Then, I distribute point values for exam questions, prompts, and tasks evenly across all Bloom's levels. First, I write exam questions, then I assign each question to its Bloom's level, decide on the total points available on the exam, and distribute those points evenly across the Bloom's levels. This way, students have opportunities from the basic (factual recall in the Remember level) to the advanced (synthesizing and generating information in the Create level) to demonstrate the extent of their mastery. For example, if a student can only accomplish Bloom's level 1 tasks (Remember), e.g. "define" and "label," then they will earn between 0 and 20% points and an F grade. If they can also complete some of the Bloom's level 2 tasks (Understand), then they move into the D letter grade range (earning 20-40% of available points).



This approach, which I call "Bloom's Grading," creates at least two powerful outcomes. First, by aligning letter grades with Bloom's taxonomy, which is a well-described and widely-used concept, I produce a letter grade rubric. I include this rubric in my course syllabus, because it clearly defines what tasks I expect students to be able to perform for each letter grade.

Second, Bloom's Grading reduces cheating. One letter grade worth of points (20%) is available from upper Bloom's level (Create and Evaluate) questions, which are essentially cheat-proof. When a student is asked to create or to provide an evaluation of something, the instructor has every expectation that the response will be unique. Most students know to avoid this kind of cheating, because it is so easily detected. Likewise, the mid-level tasks (Analyze and Apply) are cheat-resistant, depending on how you word questions and prompts. I often append a "Briefly explain your response" requirement to elicit a unique response from each student. These higher Bloom's questions also resist cheating because they usually don't have responses that are available by a search engine query. However, that's not necessarily true of the lower Bloom's questions, which are often delivered as define, label, true/false, multiple-choice, or fill-in-the-blank type questions.

Importantly, because of the even distribution of points across Bloom's levels (and thus letter grades) in this approach, even if students dishonestly acquire answers to your exam questions, that still won't earn them a good grade. For example, if they succeed at cheating by looking up the answers to all of your Bloom's level 1 and 2 questions, and if they earn full points on all of them (40% of total points), they would only have earned a D letter grade by their cheating. Moreover, cheating takes time! On a timed exam, students can't afford to look up the answers to some or most questions, because they won't have enough time to address the 40-60% of available points on your cheat-resistant and cheat-proof upper Bloom's level questions.

This grading approach is not ideal in all circumstances: it has flaws and drawbacks. In any online exam, students can cheat by communicating answers with each other. Higher-level Bloom's questions still resist this weakness, but direct messaging can provide organized groups of students a fast method of cheating on lower-level questions without resorting to web searching. Also, as always, each instructor has to decide how much effort they're willing to devote to enforcing academic honesty. There is no equitable way to make any exam entirely cheat-proof, especially online, and there is a direct trade-off between how cheatable an exam is and how much time an instructor invests in creating the questions and grading the responses. Let's reduce the efficacy of cheating, and thus the temptation to cheat, by using Bloom's Grading to actively and thoughtfully increase the relative weight of higher Bloom's tasks.

#### References

- "A taxonomy for learning, teaching, and assessing: a revision of Bloom's Taxonomy of Educational Objectives." Anderson and Krathwohl, eds. New York: Longman. 2001.
- Durm (1993) "An A is not an A is not an A: a history of grading" *Educ Forum* 57.