

What are the principles of effective grading?

Grading is not usually an aspect of teaching that we look forward to. It's hard work that requires we read, respond to, and score numerous assignments, and then decide what those scores mean. Not only is it work-intensive, it also feels very different from other aspects of teaching: rather than the energetic give and take that happens in our classrooms, grading feels like a solitary activity that only impacts students insofar as they will respond with satisfaction or dissatisfaction to our evaluation of their work. But grading doesn't have to feel this way because, in fact, grading is a dynamic process that provides meaningful information to students about their development in our disciplines and how they can improve their learning. Done the right way, grading can motivate our students to work hard and persist in our courses.

How can grading motivate student hard work and persistence?

Principles for a motivating approach to grading

- Grading should inform students about their progress toward course goals.
- Grading should use a weight system that communicates to students that learning is the result of skill building.
- Grading practices should acknowledge the limitations of our assessments in productive ways.

Putting these principles into action

- **Grade in reference to set criteria rather than using a curve.** Student scores are what we end up converting into grades. Those scores can be derived from a rubric, a percentage of correct answers, a checklist of completed work, or other criteria that reflect the goals for learning that you've set for your students. In each of these examples, students' grades are determined by scores related to set standards for achievement. Grading in this way, in relation to set criteria, helps students see where they stand in relation to the learning goals that have been established for all students. When students see such criterion-referencing grades along with information about both the successful and the unsuccessful aspects of their attempts (by referring to a rubric or other criteria), they understand that each student is being measured using the same benchmarks. These criterion-referencing grades help students consider where they struggled and how to improve their performance. When students see a grade as information about where they stand now and how they can improve, the result is a stronger sense that they can develop new skills and that we have designed our assessments to help them see where they are and where they can go. This will motivate students.

Consider what happens when we do not grade in reference to set criteria and grade on a curve. This means that the instructor looks at the spread of student scores on an assessment, identifies the average score, and assigns scores based on the spread away from that average. The aim when instructors curve grades in this way is to ensure that only the highest 2% of students receive A's, the next 4% of high scoring students get B's, the next 68% of students get C's, and the lower tail of scores result in 4% of students

receiving D's and the final lowest 2% receiving F's. In effect what we are doing is allowing all students' scores to determine the meaning of any individual student's score. What a demotivating message we send to students: grading in this way tells them that their individual efforts at learning count only insofar as how they measure up to others' efforts. Students in this situation feel they have little control over their learning, may be less willing to work cooperatively with their peers (as they now must compete with them in order to attain good grades), and will focus less on learning and more on performing for a grade (Ambrose, Bridges, Lovett, DiPietro, & Norman, 2010). Ultimately we must ask ourselves, why we want to see a normal curve in student learning: while the shoe sizes and heights of our students should follow a normal curve, as instructors we should plan and teach our courses to ensure that most of our students achieve our course goals.

If we want students to use grades to more fully understand their learning and act on that understanding, grading in relation to set criteria is the only reasonable way to grade.

- **Use weighting systems when you plan and design how you will assess students.** Our assessment plans and our individual assessments are designed to move students toward our learning goals, and one way to signal to students how the parts of the plan work together to help them learn is to use weighting systems that motivate them by indicating how we expect skills to build and what skills really matter for their learning in our courses. There are three key ways to use weighting systems to motivate students in this way.
 - *When students make multiple attempts at learning, weight first attempts more lightly than later attempts.* The most effective assessment plans require students to make multiple attempts at the learning we want them to demonstrate. For example, rather than asking students to produce one complex paper at the end of a course, it is more productive to require students write four papers where they practice the learning we want them to achieve and get feedback to improve these iterative attempts. Rather than asking students to take two large exams, it is more productive to require students take five somewhat shorter assessments and use the feedback they get to improve their studying and problem solving skills. When multiple assessments allow students to build skills iteratively or to work on parts of a larger paper or project so that they can see how the parts connect, students are guided through skill building in a highly motivating way, and grading also can be understood as feedback on those emerging skills. For this reason it is advisable to weight the first attempts lower than later attempts when you use multiple, iterative assessments.
 - *Break assessments into parts and weight those parts according to their importance in relation to your course goals.* When we have designed a rigorous assessment, that assessment is often an authentic one that has realistic complexity and thus is a multi-part assessment. Such an assessment might be a three-part paper or an exam that requires students to first solve some problems and then explain the principles they used to solve them and the places where they are still uncertain as to their approach. When we look closely at an assessment that requires students to think in different ways and draw on different kinds of skills, we realize that some of the

learning we are asking students to demonstrate is more important than other kinds of learning that the assessment requires. For example, we may require students in a statistics course to write up their findings from a small research project they conduct, but is the writing entailed as important as the logic of their inquiry and the thinking behind their interpretation of their results? We may require students in a physics class to attempt some problems on an exam, but are their answers as important as the thinking that motivated those answers? As we design assessments, it is beneficial for us to recognize that our assessments require multiple skills and that we need to help students develop those skills. This kind of assessment analysis can be challenging, but if we note those skills required by an assessment and ask ourselves which of the skills we can really help students practice, we can then weight skills and parts of assignments in a way that indicates to us and to our students where the most practice and effort is to be productively focused.

- *Demonstrate to students that they will have multiple opportunities to develop and demonstrate their learning by communicating the weight of each assignment on your syllabus and in your assessment descriptions.* We may have taken the time to create assessment plans and assessments that are structured using the principles of skill building and iterative practice, but we must communicate to students clearly that we have weighted assignments accordingly. Be sure that your grade breakdown on your syllabus and in your assignments reflects the thought you've put into your weighting system.
- **Let error of measure help you and your students learn by reflecting on and acting on the limitations of measurement.** A fear that many instructors have is that students will question the grades that they receive. When we return graded exams, we might hear, "The wording on this question confused me!" After receiving a graded paper, a student might complain, "You never really explained how you wanted us to develop that part of the essay!" Students might take their ideas further and suggest, "You should drop that question!" or "You should give us some bonus points!" While these kinds of student comments at first seem like students making excuses for their poor performance and grade grubbing, they actually point to two important aspects of any assessment: validity and reliability. We want to believe our assessments do truly assess student achievement in our disciplines, but sometimes we design an assessment that ends up assessing what we didn't intend to assess. The first complaint above may actually indicate that a multiple choice item tapped students' reading ability and not their disciplinary thinking. We also want to believe that our assessments are a fairly true measure of student achievement, but sometimes an assessment reveals less about student learning and more about preparation or practice. The second complaint above may actually indicate that students weren't well prepared for a particular aspect of the assessment and that the assessment is not a reliable measure of their abilities.
We ask a great deal of ourselves and our assessments: essentially we ask that our assessment plan acts as a sound measure of student achievement that can render valid and reliable scores and grades. In some ways we are harder on ourselves and our assessments than we are on the measurement instruments we use every day! There is a reason we weigh ourselves multiple times or take our blood pressure more than once. We know there is error in any measuring instrument, so why do we think that our

measuring instruments (our assessments) are any different? The reason is likely that when we acknowledge error in our assessments, we begin to worry if those assessments are defensible. But there is another way to respond to the reality that our assessments are not perfect (and that we are also fallible!): we can acknowledge that error is a feature of every measure and every assessment and work with that accordingly. There are several key strategies that can help you and your students work productively with error of measure. By using these strategies, you remind your students and yourself that assessment is a learning tool used to discover what students know, not a mechanism to create grades.

- *Allow students to make multiple attempts on assessments.* If your assessments are multiple choice or short answer exams, create a pool of items that you can draw from to create parallel exams (exams that tap the same kinds of learning). Students can take a version of their first exam, get feedback that focuses their review of the material and then take the second version of the exam. Students can be graded using an average of these parallel exams, or they can choose to drop the lower score. If your assessments result in written work, the suggestions above about designing a series of iterative attempts will serve the same purpose of ensuring that students have multiple opportunities to demonstrate their learning.

- *Use item analysis to see where students struggled, ask yourself why, and communicate your findings clearly to students.* After students have completed an assessment and you've scored their efforts, take some time to look over what items or what sections of an assessment were challenging for students. If a large percentage of students scored low or failed in one area, ask yourself why. Perhaps students had little preparation or practice, the wording or task was not clear to students, the item was keyed wrong, or you yourself made some sort of grading error. In these situations it's wise to communicate to students that that item or part of your assessment didn't measure student learning well and tell them your plan for managing that measurement problem.

- *Allow students to rework problems or parts of assessments.* In the event that students underperform on an assessment, it is beneficial to have them rework that area of an assessment. If you recognize that the cause of student underperformance is that you have not given students sufficient preparation time or practice experiences, reteach and then reassess."

- *Drop the lowest grade.* If students have the opportunity to take multiple assessments in your course, your grading policy can allow students to drop the assessment with the lowest grade. This is a robust way to ensure that the one poor performance a student has or the one less reliable or valid assessment you designed does not factor into a final grade.

- *Allow students to addend an assessment.* When you design assessments, your aim is to get a clear picture of student learning, but assessments must be manageable and for that reason exams and papers sometimes can't capture everything a student has learned. Because any assessment requires them to focus on some skills but not necessarily on all the skills learned in a course, students sometimes feel that something they understand deeply or new conceptual or procedural abilities are not tapped on an assessment and the result is that they feel frustrated that the assessment didn't allow them to demonstrate their learning fully. To ease this kind of frustration, allow students to respond to this prompt: "Now that you've completed this assessment, you may

realize that you've learned other things in this section of our course that you'd like to show me. Take some time to explain in your own words two concepts/skills/processes you've learned, give a novel example of each, and describe how you will use this new knowledge in the future (in this class, in another class, or in your life)." Give students credit for well-developed assessment addenda.

○ *Don't grade things that are impossible to assess.* We sometimes think that we will motivate students by giving them points for effort or for participation. But how can we actually measure dispositions and behaviors? We can't unless we end up tallying or counting observations like how many times a student speaks in class. This can lead to meaningless behaviors on our students' part and time-consuming bean counting on our part. But there is a larger problem with grading dispositions and behaviors—they reveal nothing about student learning. For students to learn, they need information about their thinking, but throwing points at students for simply being present in our classes or appearing enthusiastic does not supply that information. When 10% or even 20% of a grade is meant to reflect student dispositions and behaviors, we cheat students of real feedback on their learning and introduce a large amount of error into the measure of their achievement.

○ *Grade papers and other complex projects when you have energy for the task.* A final suggestion is to remember that your disciplinary expertise is a key part of how you evaluate and grade student work. If you are feeling tired or out-of-sorts, your energy and mood may make it difficult for you to read or respond to student work as clearly as you would if you tackle grading in a fresh frame of mind. If you are requiring students to create papers or projects, having a well-developed, valid and reliable rubric helps you focus your thinking about the work you respond to. But grading must be paced so that you can use that rubric and your expertise well and so that fatigue doesn't introduce more error into your grading. When it is time to grade student work, do so in manageable spans of time, with breaks in between. Grade a little each day so that you don't feel overwhelmed and so that you can approach student work with optimism and an open mind.

Examples of grading systems that motivate students

A grading system for course assessments that helps students see that a course has been designed to provide opportunities for learning

This is an example of the grading system in a public policy course. Of note here is that the grade breakdown for the course shows students how many opportunities they have to practice skills and build them.

Components of final grade. You will work on 3 public policy memos this semester. I have created a cycle of steps that will prepare you well so that you can produce a policy memo that is well researched and has received feedback from two peers and from me. We will have regular homework that will help you make the most of the course readings. I have also created three assignments that allow you to reflect on your progress and on your peers' development as peer editors. Note that I will allow you to completely revise your approach to one of your three policy memos.

- 5% Public policy memo 1 preparation

- 5% Public policy memo 1 draft
- 3% Public policy memo 1 peer feedback
- 7% Public policy memo 1 final draft

- 5% Public policy memo 2 preparation
- 5% Public policy memo 2 draft
- 3% Public policy memo 2 peer feedback
- 7% Public policy memo 2 final draft

- 5% Public policy memo 3 preparation
- 5% Public policy memo 3 draft
- 3% Public policy memo 3 peer feedback
- 7% Public policy memo 3 final draft

- 5% Public policy memo writing reflection midterm
- 5% Public policy writing reflection final
- 2% each 10 reading response homework challenges (20%)
- 3% peer evaluation
- 7% Annotated approach revision of one of your three public policy memos

A weighting system for a written assessment that helps students see what skills are to be their focus as they work on an assessment

This is an example of a weighting system used to create a grade for a psychology case study paper. This is not the rubric that would be used for this assignment, but simply shows the percentage value assigned to each section of the rubric.

- 15% Use of psychology concepts in problem analysis section
- 15% Use of psychology concepts in solution analysis section
- 10% Use of concrete details from case to establish problem analysis
- 10% Use of concrete details from case to establish solution proposal
- 15% Use of psychological research and concepts in evaluation of solution
- 5% Annotated reference #1
- 5% Annotated reference #2
- 5% Annotated reference #3
- 5% Annotated reference #4
- 10% Organization and logic of thinking
- 5% Use of written conventions

Grading policies that acknowledge and work productively with error of measure
Here is an example from a biology syllabus.

A note about exams. Along with other assessments, I have written five short exams to assess your learning this semester. The exams come at the end of a sequence to help me and you see what you can do and where you are still struggling. The exam questions are

mostly multiple choice, but they are challenging and require you to apply course concepts and analyze realistic problems using those concepts. An exam is not a perfect measure of your learning, however, so I have put in place some policies with that in mind.

1. Your first exam is worth less than your subsequent exams because you may need to build up your studying and preparation skills.
2. You can “rewrite” up to three incorrect exam questions on each exam. This means that you will explain why you came up with your answer, why the right answer is best, what part of your thinking is now clarified from having made your error, and how you will use this new learning moving forward (give a concrete example).
3. On each exam you will find the same open-ended question at the end of the exam: “In this space you can demonstrate what you know about a biology concept we worked with during this course sequence that my questions didn’t get at. Explain the concept in your own words, describe why it is important by providing an example of it, and articulate how this concept has developed your thinking about biology. Be detailed and concrete in your responses.
4. If I find that 60% or more of students have missed a test question, I will review concepts related to that item with the class and drop it from the test grade.

Resources about effective grading

- Ambrose, S. A., Bridges, M. W., Lovett, M. C., DiPietro, M., & Norman, M. K. (2010). What factors motivate students to learn? In *How learning works* (pp. 66-90). Jossey-Bass.
- Nilson, L. B. (2016). *Teaching at its best: A research-based resource for college instructors* (4th ed.). Jossey-Bass.
- Walvoord, B. E., & Anderson, V. J. (2010). *Effective grading: A tool for learning and assessment* (2nd ed.). Jossey-Bass.

If you’d like support as you consider your approach to grading or a grading scheme, please feel free to [request a consultation](#).