GRADING: DO IT LESS, DO IT BETTER   Peter Elbow

PROBLEMS WITH REGULAR GRADES
ÏGrades don't have clear meaning: how good or bad is "B" or "89"?
ÏGrades give no feedback about *what* was good or bad--learned and not learned.
ÏA one-dimensional number (or letter) cannot validly represent the quality of a multidimensional
     performance such as a piece of writing or a student's work over a semester. (Some parts of the
     writing or course work will be better than others. Different graders will give different grades
     according to how they value the different dimensions.)
ÏGrading undermines the teaching-and-learning climate. It leads many students:
   --to work more for the sake of the grade than for learning;
   --to resent their grade and resent us. This leads to an adversarial climate where they feel us as the
     enemy rather than ally in the learning process.
ÏGrading is difficult and often makes us anxious because we know that true fairness is impossible.

DO IT LESS
Ungraded assignments:
There's no law against *requiring* work, but not *grading* it. It actually makes good sense. (If necessary,
grade on a binary scale: 100% for doing it; 0% for not doing it.)
Examples of ungraded assignments:
   --Low-stakes writing like freewriting, journal writing, and letters.
   --Starting the semester with an ungraded week or two: lots of required work, all ungraded.
   --Practice tests and exercises.

Portfolios:
ÏA grade on an individual piece of work is much less trustworthy than a grade on a body of work.
ÏPortfolios invite the grade to reflect students' abilities at the end of the course--rather than an
     average pulled down by how bad they were before they learned what you were teaching.
ÏIt's not useful to keep students entirely in the dark about their grade for the whole semester. It
     helps to have a midsemester portfolio with a provisional grade. (Some grade individual essays,
     but that defeats one of the main advantages of portfolios: saving the time and headache of
     assigning (untrustworthy) grades to individual pieces.)

Minimal Grading:
ÏUse fewer levels of quality; settle for cruder differentiations. For example: excellent, good, fair,
     fail; 1,2,3,4;   strong, satisfactory, weak;   pass/fail.
ÏTeachers sometimes resist the idea of fewer levels because they assume that only low stakes work
     can be graded with few levels (or even pass/fail)--and that high stakes work must be evaluated
     with a full range of levels. But this assumption confuses *stakes* and *levels*. High *stakes* makes
     students work harder; many *levels* of quality just makes *us* work harder. There is no harm in
     using fewer levels for high stakes work: students will work hard if we simply raise the bar high
     enough for "satisfactory" or "good"--or even for "pass."
ÏStudents sometimes resist the idea of fewer levels because they have an instinctive fear of losing
     information--losing the greater evaluative information that comes from fine distinctions--even
     when, really, they know that information is untrustworthy. This shows how anxious and
     distressed people are about being evaluated--how difficult it is to think clearly in this area.
Ï"But how can I calculate a final grade with twelve levels if all my intermediate grades have many
     fewer levels?" In fact there are many possible systems that are rational and easy to work out.
ÏSummary claim: minimal grading is easier, fairer, more accurate, and does less harm to the climate
     for teaching and learning.

Contracts. See the handout. I'll talk about contract grading at the end.

DO IT BETTER

Explicit Criteria.

Ÿ When we assign work, it helps to announce the criteria by which it will be judged. This will help students do better work--for both cognitive and emotional reasons.

Ÿ Sometimes, when we are reading their work, we are struck by an important criterion or two we *should* have announced. Next time we'll do better.

Ÿ Using explicit criteria makes our grading fairer: it keeps us from letting our response to the entire work be too swayed because one dimension is particularly good or awful.

Grids. They are quantitative, but they give substantive feedback about what was good or bad, learned or not learned. An example:

Weak	Satisfactory	Strong	
☐	☐	☐	Genuine revision, substantive changes, not just editing
☐	☐	☐	Ideas, insights, thinking
☐	☐	☐	Organization, structure, guiding the reader
☐	☐	☐	Language, sentences, wording
☐	☐	☐	Mechanics: spelling, grammar, punctuation, proofreading
☐	☐	☐	Overall

Example of a grid used as a mini-file on a computer. Import it into each individual evaluation letter, and then write comments after particular criteria as needed:

> **Genuine Revision, Substantive Changes, Not Just Editing**:
>
> **Ideas, Insights, Thinking:** *Strong. I liked the way you complicated things by exploring points that conflict with your main point.*
>
> **Organization, Structure, Guiding the Reader:** *Weak. I kept feeling confused about where you were going--though also sensing that my confusion came from your process of complicating your thinking. This confusion would be GOOD if it weren't a final draft.*
>
> **Language, Sentences, Wording:**
>
> **Mechanics: Spelling, Grammar, Punctuation, Proofreading:** *Weak. Because of all the mistakes, this paper doesn't fulfill the contract and is not acceptable. I'll call it acceptable this first time IF you give me a fully cleaned up version by next class.*
>
> **Overall**: *Unsatisfactory for now.*

I often write an additional discursive comment at the end. Grids can be used for work other than writing. We just have to figure out the criteria that matter. Even on a math exam, it is possible to differentiate the different skills required by different questions (or parts of questions).

Student self-assessment.

We can require it; it takes no work from us. Students know more than we do about many aspects of what they have learned and not learned--certainly about how much work they did and about what kinds of learning processes they went through. (See the example with the contract supplement.)

SUMMING UP.

Regular grades are all number and no meaning: all vertical (levels) and no criteria (horizontal). Minimal grading gives us less of the vertical. Criteria and grids give us more of the horizontal.

Contract grading works best for me. It's radically minimal but maximizes information.

Ÿ Contracts reduce tension and arguments around grades. They give students control over their grade; they know pretty well at every moment what course grade they are in line for.

Ÿ With normal grading, students don't really think about my evaluative feedback; they just do what I say or imply. With the contract, they are safe for a B even if they *don't* have to do what I say. But they have to make substantive revisions. They actually have to think about whether my evaluation makes sense. Thus contracts make evaluation more effective for learning.

Ÿ Contracts let me put my effort into what I enjoy: figuring out which activities most reliably cause learning. They also let me ask directly for those qualities in writing that matter most to me (e.g., exploring perplexity). I can spend very little time doing what I hate and distrust: trying to *measure numerically the quality* of writing or learning.