

*This chapter suggests two ways to make the grading of writing easier, fairer, and more helpful for students: using minimal grades or fewer levels of quality, and using criteria that spell out the features of good writing that we are looking for in the assignment.*

## Grading Student Writing: Making It Simpler, Fairer, Clearer

*Peter Elbow*

I see three main problems involved in trying to grade writing. First is the plain *difficulty* for us in trying to figure out the grades. For each essay in the stack, we have to decide between A, A-, B+, B, B-, C+, and so forth. If we use the full set of grades, we are using eleven levels (thirteen if we use A+ and D-). Even if we never use *any* grades below C-, we are still having to make fine evaluative discriminations among eight levels. I am relieved to discover how many other faculty members admit to their struggles and frustration with trying to grade writing.

Second (and this helps explain why grading is hard), we know that these decisions are not trustworthy, no matter how hard we agonize. Careful research has demonstrated over and over what common sense has told us—and what our students have learned through controlled experiments of submitting the same paper to different teachers: good teachers and evaluators routinely disagree about grades—and disagree widely. Even a somewhat conservative defender of grades like Edward White (1990) warns of the danger of believing there is such a thing as a “true grade” or “right grade” for a piece of writing.

Third, grading tends to undermine the climate for teaching and learning. Once we start grading their work, students are tempted to study or work for the grade rather than for learning. They see writing as an exercise in trying to say “what teachers want” rather than working out their own thinking. Students resent the grades we give or haggle over them and, in general, see us as people they have to deceive and hide from rather than people they want to take into their confidence.

I don't see any way we can fully *eliminate* these problems if we grade at all (and I am not proposing to give up grading altogether—in this chapter). But I

think I can suggest two major ways to *lessen* these problems: minimal grading and giving grades more meaning.

## Minimal Grading

**Using Minimal Grades on Low Stakes Writing.** Minimal grading has always been around, but it is much more common now because so many teachers in so many disciplines use low stakes writing: frequent informal assignments designed to get students to reflect on what they are learning from discussions, readings, lectures, and their own thinking—assignments that count for credit but that individually, don't bear *heavily* on the final course grade.

Thus minimal grading is often the result of crude pragmatic pressures. These faculty who use writing for learning are often not writing teachers, and they are often teaching large classes that are not writing courses. As a result, these faculty often feel they cannot grade student writing carefully or elaborately—and they don't feel obliged to do so because the pieces of writing don't count so much. In this situation, these faculty members have turned to a variety of forms of minimal grading:

A scale with three levels: for example, strong/satisfactory/weak or excellent/OK/no credit.

A scale with two levels: for example, pass/fail, satisfactory/no credit, or check/minus.

A "scale" with only one level: the assignment is acceptable if it is simply turned in at all.

A zero scale: faculty members sometimes don't even collect some pieces of writing-to-learn. On these occasions, they simply ask students to freewrite at the beginning of class about the homework reading—perhaps in relation to a topic for the day—or to write at the end of the class about the day's discussion or lecture, or to write in the middle of a class about an issue that has come up, especially when a discussion goes dead.

Faculty members often don't make any comments at all on this writing. Thus the most obvious advantage of minimal grading is simplicity. It's much less onerous to read lots of student writing when the grade is quick and easy to give and we don't have to comment.

But there are also advantages for student learning. Even though minimal grading removes the incentive to strive for an A for excellence (though we get a fair amount of this incentive if we use a three-level scale), we get to ask students to write far *more* than if we had to grade everything carefully. We get to ask them to think actively about far more of the course material. They have to answer the questions, get their thoughts into writing—yet they don't have to worry so much about whether they are writing in the way that the teacher likes or saying what the teacher agrees with.

Some people complain that minimal grading takes away motivation, but when students struggle for excellence *only* for the sake of a grade, what we see is not motivation but the atrophy of motivation: the gradual decline of the ability to work or think or wonder under one's own steam. Minimal grading on low stakes assignments, however, is a way to help students gradually develop a bit of *intrinsic* motivation—develop a bit of their own curiosity and standards. They get a time out from their habitual and understandable preoccupation with “What is the teacher looking for?” They get a chance to ask themselves, “What am *I* looking for? What do I think? What are my standards?” Of course, students nurtured in a grading economy often need *some* extrinsic motivation to get them working. But that's exactly what minimal grading provides. It makes them *do* the writing and engage the material, but it gives them a lot of choice about *how*. Thus they get small protected spaces for gradually developing small bits of intrinsic motivation. And of course, they still have some high stakes assignments that we grade in a high stakes way—assignments where we provide most of the motivation. (I have written more in Chapter One about the advantages of low stakes writing.)

A few students are confused at first by this dialectic between low stakes and high stakes grading. Despite my explanations and warnings, they are caught off balance the first time I use high stakes grading on a high stakes essay. They assume that because I started off with low stakes grading for a number of assignments, I am a “low stakes kind of a guy”—for students often pigeonhole teachers as *either* high *or* low stakes in their approach, as either “hard” or “soft.” I find I have to be extremely explicit and even repetitive to help them realize that this whole approach enacts a dialectic. Indeed, I like to bring in the word *dialectic* and talk about the psychological and intellectual benefits of unresolved contraries. But I have learned also to be more blunt and say, “Perhaps you better think of my grading policy as *schizophrenic*.”

Before concluding this section, let me make a point of strategy about *introducing* minimal grading on low stakes writing at the beginning of a course. I always start with a two-level scheme such as pass/fail or satisfactory/No credit. I don't start with the three-level scale because I don't want students to become preoccupied with getting that grade of strong or excellent—preoccupied with trying to figure out what I am looking for; I don't want them to fall back into writing for the grade instead of writing for exploration or learning. And I don't start with no grading at all (accepting anything they turn in or not even collecting the writing at all) because I need to exert *some* pressure in the beginning to engage or focus their efforts. Otherwise, students who have never done this kind of writing are too tempted not to try at all. In short, I start with two-level grading because I want to teach an important skill that interestingly is somewhat rare in students: how to write about an academic topic with their mind focused *wholly* on the topic and issues—and not on how the writing will be graded. Without that skill, it is hard for them to experience satisfaction or even pleasure in writing seriously about something they are studying.

After students have done three to five pieces of low stakes writing graded on a two-level scale—that is, with a great deal of leeway but still a threat of no credit for goofing off—I can move my minimal grading in either direction if it suits me. I can do some low stakes writing in class that I don't even collect—and trust that they will make good academic use of the time. And I can sometimes use a three-level scale to push some of them harder with the reward of excellence—and trust that they won't just clench or become too preoccupied with what the teacher wants and style. But in fact, I often stick with the simplicity of two-level grading throughout the semester.

**Using Minimal Grades on High Stakes Writing.** We tend to link minimal grading with low stakes writing. That is, we tend to assume that if we want to raise the stakes of the writing and make students strive harder, we should grade on more levels. Conventional A through F grades *feel* more serious.

But this assumption is mistaken. After all, some teachers *do* grade low stakes assignments with the eleven levels of conventional grades. More importantly, we *can* grade high stakes writing assignments with minimal levels. For example, at M.I.T. for the last twenty years or so, faculty have given nothing but pass and fail as final grades to all first-year students in all courses. The stakes are very high indeed and so are the standards, but only two levels are used.

The important strategic point here is that stakes and levels are two quite different dimensions of grading. We will have a much better time grading high stakes assignments if we decouple them—sever the unconscious link between them—and realize that grading always involves two very different questions: How much credit is at stake in this performance? How many levels or grades shall I use on my evaluation scale?

It is true that if we add more levels we add more opportunities to record excellence or distinction. But this also means more competition, more hierarchy, more pecking order—fewer people in same boat. My point here is that increasing the number of levels and giving students the chance to get an A rather than a pass doesn't always succeed in making students work harder:

- The stakes have to be high enough to make extra effort worth expending. Why work hard for an A if it has very little effect on your final grade? A few students work hard for a low stakes A, but usually it's only because they are interested or eager to learn—not merely because of the chance to get an A rather than a pass.
- A fair number of students don't feel an A is within their reach. Sometimes these students will work harder for a strong or even for a demanding pass than for the B-'s and C+'s they are accustomed to getting—grades that they experience as a put-down.
- Some students are simply turned off or thrown off stride by the competition itself.
- Finally, there are students who *can* get A's, but they don't care enough to work for them. Many students settle for doing mediocre work

because of their feelings about the course or the conditions in their lives. It's amazing what good work most students can do when they really work hard under supportive conditions. (There are many cases of students wrongly accused of plagiarism because their teachers didn't think they were capable of such good work.)

In short, if our goal is to get students to work harder and more carefully on certain pieces of writing, increasing the number of levels in our grading scale is a very uncertain engine for getting there. But it is a *certain* engine for making us work harder. More levels mean more discriminations to make—discriminations that by the same token are harder to make because they are more fine grained and therefore more debatable. In making these finer distinctions, we are providing students with more occasions to dispute and resent the very grades we struggled so hard to determine (“What do you mean B minus? That was a solid B!”).

We have a far better chance of getting students to work hard and carefully if we raise the stakes—and then think more strategically about raising the number of levels only a little bit or even not at all. This means not increasing our workload so much and not giving students so many opportunities to resent or quarrel with our grades. If we pursue this approach, we have two important variables we can manipulate.

1. We can make a strategic choice about how many levels to use—even within minimal grading. Two levels is easiest for us, but it doesn't give much scope or incentive. If we move to three levels (for example, strong/satisfactory/weak), we give much more incentive—yet increase our work only a little. Some teachers use four levels (for example, poor/fair/good/excellent). This gives them a lot more work than three levels and provides students more occasions for disagreement or resentment—and yet even four levels is easier than conventional grading.

2. We can make a strategic choice about where to put the borderline between levels—about, that is, how high we set the bar. When I wrote in the previous paragraph that a two-level scale didn't “give much scope or incentive,” I was sliding along with a conventional easygoing assumption that “pass/fail means no sweat.” This assumption comes from low stakes writing. But there is no law that passing has to be easy—especially for high stakes writing. A higher threshold or demand can be natural and appropriate. Note that even a two-level scale can be very demanding if we raise the bar (as at M.I.T.). If we use three levels, we have even more scope for making strategic decisions about where to place the bars.

Notice, by the way, that with low stakes assignments we can get away with quietly slipping the bar up and down a bit depending on the student. That is, we can tacitly insist that skilled and well-prepared students do a better job for satisfactory than unskilled students with a weak background who are trying hard and improving. When the stakes are low, students aren't likely to mind or even notice a little bit of this kind of flexibility—as long as we make it

possible for all students to get a satisfactory without inordinate effort. I at least don't want to push students so hard on low stakes assignments that I lose the essential benefit: that they must engage the material seriously but still get to risk and explore—and that we don't have to work too hard in applying the standards.

In this section, then, I'm suggesting that even though we associate minimal grading with low stakes assignments, it is also appropriate and helpful for high stakes assignments. We have the option of raising the stakes and yet keeping down the number of levels in our grading. Such pragmatic considerations are particularly important for large classes. If we are looking at a big stack of papers, think about the difference between giving each one a conventional grade and just picking out the ones that are notably strong or notably weak (that is, using a three-level scale).

**Minimal Grading and the Final Grade.** At this point, some readers will naturally ask, "But how can I calculate a conventional grade for the course if I only have minimal grades to work with?" This problem brings up another important variable in grading: the number of assignments we grade. If we only have two or three graded assignments and they are graded on only two or three levels, then we have no basis for calculating the final grade for the course. But if we have *lots* of minimal grades—which is easy and natural with lots of low stakes assignments—then it is no problem to derive a conventional final grade.

Take an extreme example. Suppose we have a course with weekly low stakes writing assignments judged on a two-level scale, two high stakes essays judged on a three-level scale, and a conventional final exam judged on a three- or four-level scale. (It's not so hard or unreliable to use four levels for an exam if the exam contains multiple answers.) We can easily calculate a final course grade as follows: students who have a satisfactory on all the low stakes pieces start off with a foundation of B. Then the two high stakes essays and the exam determine whether their B gets pulled up or down. Students who don't have consistently satisfactory work on their weekly low stakes pieces start off from a lower foundation, and they are similarly pulled up or down by their high stakes work.

In this formula, I give quite a lot of emphasis to the low stakes assignments: individually, their weight is low, but in sum, they count for a lot. I want my students to take these pieces seriously—that is, to work hard—but not in a worrying way. Not everyone will agree with my priorities here, but there are obviously other simple formulas one could use for this situation. I just want to illustrate that we can easily derive a maximum-level final grade from minimum-level constituent grades. And I didn't even talk about other factors that many of us feel are important ingredients in calculating a final grade: attendance, participation, effort, and improvement.

*Portfolios.* Portfolios can be a big help in deriving a conventional final grade from a pool of minimal grades. To decide whether an individual piece of writing is a B or a B- is, in my view, to produce a worthless decision. To decide whether a *portfolio* of writing is a B or a B- is more justifiable: with fuller and

richer data, we are justified in trying for a fuller richer discrimination. (And yet I still wouldn't try for such fine discriminations in grading a portfolio—and in fact it turns out that professional readers of portfolios tend to disagree with each other in their grading as much as readers of individual papers. I think it makes more sense to use minimal grading on portfolios too, and then let this minimal yet high stakes portfolio grade be one among various factors in calculating a final conventional grade.)

The real value of portfolio grading is not for precision but for enhancing learning. Portfolios give us a way to get students to look back over all their work for the course and reflect carefully on what they have learned—and even to analyze their learning process. For the most important piece in a portfolio is the reflective essay on the contents of the portfolio.

Portfolios are probably not so feasible for very large classes. You can't even carry one hundred portfolios back to your office. But in a large class, it makes perfect sense to ask students to *keep a portfolio* and then at the end to look back at it analytically and write a reflective or analytic essay about their learning and writing. (Some teachers even ask students to suggest a final grade in this final essay of self-analysis.) For low stakes assignments create much more learning if we ask students to gather them all together at the end and work out the most important insights they can—both about the course material and about their processes of learning, thinking, and writing. (See Fulwiler's suggestion in Chapter Two for a portfolio of letters written throughout the course.) Low stakes writing provides students a particularly good window for reflecting on their own habits of thinking and using language. And I love the way portfolios permit me to invite students to put some of their low stakes writing into the final high stakes portfolio—a satisfying thing to do because some of their low stakes writing is often very good. (Readers who want to open the door into the rich literature on portfolios can start with the Belanoff and Dickson, 1991, and Yancey, 1992, collections.)

*Contracts for Grading.* A contract makes the final grade easy—indeed almost automatic—even if one uses minimal grades during the semester. A contract says, in effect, “If you do X, Y, and Z, you can count on such and such a grade.” The point of a contract is to focus less on trying to measure degrees of quality of writing and instead to emphasize *activities* and *behaviors* that will lead to learning.<sup>1</sup>

Teachers tend to think of grading as a way to produce behavior or even motivation for behavior, but notice how it is an *indirect* way of doing so. That is, we hope that the awarding of fair grades will cause students to engage in the learning activities we want them to engage in. Why not instead be more direct with a contract and just, as it were, *make* them do the things that we think will lead to learning? I would rather put my effort into trying to figure out which activities will lead to learning than into trying to measure the exact quality of the final product students turn in.

**Objections to Minimal Grading.** “We already use minimal grading: most faculty already give nothing but A's and B's.” I would reply that it is exactly this

restricted use of the grading scale that has exacerbated two intertwined problems: grade inflation and grade meaninglessness. When some faculty members give a full range of grades and others give mostly A's and B's, we have a situation of semantic chaos. The grade of B has become particularly ambiguous. For some faculty members, it means "good strong competent work"—and they point out that B is an official honors grade in most college and university catalogues. But when other faculty members give a B they mean "disappointing, second-rate work." And students tend to interpret B as unsatisfactory. If instead of A and B, teachers would use honest words like excellent, honors, outstanding, strong, satisfactory, weak, poor, and unsatisfactory, all parties to grading would have a better understanding of the message.

*"But minimal grading doesn't really solve this problem. Grades are just as ambiguous if most students get a satisfactory."* Not so. Even if most students get a satisfactory and thus the grade is given to a wide level of performances (which will tend to happen), the result is *not* so ambiguous as the present grading system. For we will have *clearly communicated* that the grade has this wide range of meanings through the use of an honest word (satisfactory) and because we have used a limited scale. We won't have the ambiguity of the present situation where no one knows whether B means satisfactory or not satisfactory or whether B is being used for a wide or a restricted range of performances.

*"Some teachers will probably still give mostly excellent or strong."* This need not be seen as a problem, because teachers have, as it were, signed their names to those *words*: they are certifying that the majority of performances in the class are in fact excellent or strong or honors level. Surely it can happen that most of the performances on a paper or even in a course are in truth excellent or strong—and we want to sign our name honestly to that report. In the present situation, when a teacher gives mostly A-'s and B+'s, no one knows whether she is saying, "This was a remarkable outcome," or just, "It's not too hard to get a B+ in my system." Besides, the worst grade inflation is not at the upper level but the lower level. Most teachers give passing grades and even C's to performances that they do not consider satisfactory. There would be less of this if they had to use the word satisfactory.

### **Giving Grades More Meaning: Using Explicit Criteria**

I have been arguing that we can get students to work hard and invest themselves in certain assignments if we raise the stakes but still use minimal grading—meanwhile making strategic decisions about how high to raise the stakes, how many minimal levels to use, and where to place the bars or thresholds between levels. We won't have to struggle with the eleven levels of conventional grading, and the results will be fairer. Why struggle to distinguish between a B and a B-, giving ourselves more work and our students more occasions for resentment, when the resulting decision lacks not just meaning but fairness.

Nevertheless, if the stakes are high there *will* be anxiety and potential argument around where we set those borderlines between satisfactory and



unsatisfactory, or between satisfactory and excellent. So how do we decide how high to place the bar? How do we figure out exactly what we mean by “a satisfactory essay” or “an excellent essay”? This is where we get into head scratching with ourselves, honest disagreement with colleagues, and unpleasant arguments with students.

We can never make grading completely easy or completely fair—but in this second section, I think I can show a path toward improvement. The key is to think about the *information* or *meaning* carried by a grade.

Despite the advantages of minimal grades, we must admit that conventional grades carry *more* information. That is, conventional grades are more precise than minimal grades at telling *how well* or *how badly* students did at a task. However, this added information is not only untrustworthy; it is empty. That is, conventional grades tell nothing at all about *what it is* that the student did well or badly; the greater precision of conventional grades is utterly untrustworthy. In fact, we don’t even *get* that alleged precision unless we see the range of grades for the whole class. But however bad the added information is, the fact remains that conventional grades sort students into more groups that are more finely differentiated, giving students a sense of seeing themselves as better and worse in relation to *more* of their peers. In short, however dubious the value or reliability of conventional grading, students tend at first to experience minimal grading as *taking something away from them*—and something deeply valued.

My purpose in this section is to show that we can give students *better* information or meaning in return for the bad information we take away from them. We can make minimal grades more meaningful than conventional grades if we can tell students what they are actually weak, satisfactory, or excellent *at*—and also show them that our minimal grades are actually fairer. To do so, we need to work out the *criteria* for our minimal grades.

For low stakes assignments, criteria don’t matter so much. Still, it’s not hard to come up with workable criteria that are easy to apply—so that papers can be graded quickly and with no head scratching—and in the case of large classes can even be perfunctorily checked. Here is what a typical handout might say about the criteria for weekly *thinkpieces*.

*To be Acceptable the piece must be at least 750 words. You don’t have to have a unified essay with a single thesis or point. And you don’t have to be right in everything you say about the course material. I invite you to speculate and pursue hunches. But you must seriously wrestle with or engage the academic material in the week’s reading and the topic or issue that I specify. Informal, colloquial writing is fine, but it must be clear to me as reader. Handwriting is acceptable—even a few scratch-outs and write-ins are fine—as long as the piece is neat enough to make it genuinely easy to read.*

One could add other features to the criteria: for example, that students quote a passage from the reading and work with the quote or that students be more or less right in what they say about the course material—or at least not

badly wrong. (This last criterion would require more careful reading by the teacher.)

With high stakes assignments, however, we see the greatest need and the greatest opportunity to make grades carry more genuinely useful information or meaning. If we specify and use criteria for high stakes minimal grading, we will vastly reduce uncertainty for us, and resentment, hostility, or discouragement for students.

**Using Criteria But Not Grading Each One.** Figuring out our criteria and spelling them out publicly doesn't mean we have to give a grade on each criterion on each paper. It's when we *don't* spell out criteria publicly that we have the most obligation to spell out individual reasons on each paper why it got its grade. One of our options is to tell students ahead of time what we are looking for in the papers and then give only a single minimal grade. We need this option particularly when we have a large class that doesn't center on writing and we have little or no help in teaching, or when we want to assign a lot of papers.

There is a traditional and crude distinction between *form* and *content* that many teachers use quite successfully (despite some criticism of it as old fashioned or even theoretically suspect). For example, one might explain one's criteria for a high stakes essay in a large course as follows:

*I will grade these important essays on a three-level scale, unsatisfactory, satisfactory, excellent. I will count roughly two-thirds for content and one-third for form. By content, I mean thinking, analysis, support, examples [or one might talk about specific concepts or issues in the topic]. By form, I mean clarity and correctness.*

Teachers sometimes break out these two broad criteria into four more explicit ones: correct understanding of course material, good ideas and interesting thinking, clarity, mechanics.

These are traditional but sturdy, workable criteria. Yet of course, we can work out our own criteria according to our own tastes—perhaps changing them on different papers. (Here are just a few of the diverse criteria I have seen faculty members use in grading: analysis of quantitative data, persuasion, researching new information, accuracy with sources, applying course concepts to new situations, effective revision from an early draft, effective or clear organization, establishing an appropriate relationship with readers, clear and lively voice, correct use of citation conventions, copyediting.)<sup>2</sup>

When we spell out our criteria in public—in an announcement or on a handout—we are making our grades carry more information or meaning than they usually do, even if we give nothing but a minimal grade. All too often, grading criteria are left tacit and mysterious. Also, when we spell out criteria in public, we usually grade more fairly. That is, when we lay out our criteria, we are not so likely to be unduly swayed if one particular feature of the writing is terribly weak or strong. (Research shows that teachers tend to get annoyed by papers that are full of grammar and spelling mistakes and non-

standard dialect, and consequently overlook virtues in content or reasoning in such papers.)

**Giving Grades on Individual Criteria: Using a Grid.** If we don't have too many students, we may discover that we *can* tell students how they did on each criterion. The principle of minimal grading comes to our rescue here. For just as it isn't so hard to read through a set of papers and merely pick out ones that are *notably weak* or *notably strong*, so it isn't so hard merely to note if an essay is notably weak or strong on the criteria we have named as important. Thus we might use a kind of grid and the "grade" on the high stakes paper might look like this (using traditional criteria):

CORRECT UNDERSTANDING OF COURSE MATERIAL: Excellent  
 GOOD IDEAS AND INTERESTING THINKING: Satisfactory  
 CLARITY: Satisfactory  
 MECHANICS: Satisfactory  
 OVERALL: Excellent

Notice that the overall grade is excellent but it is based on three satisfactoriness and only one excellent. That is, the use of criteria doesn't oblige us to be rigid or simpleminded in evaluating. Most teachers decide that content counts more than form. By noting strengths and weaknesses on a crude scale using multiple criteria, we give far more meaning and clarity to three-level grades than students get from the eleven levels of conventional grades. Yet doing so is usually not much harder than trying to figure out those conventional grades—and if the class is large, we can get away without writing a *verbal* comment. In truth, these crude notations on criteria are often more helpful than most of our verbal comments.

The most important advantage of this extensive use of criteria is that students at last get some substantive feedback on what they did well or badly. Most students will probably get an overall grade of satisfactory, so they will benefit enormously from knowing which dimensions of their papers were notably weak or strong. And the students with grades of unsatisfactory badly need more particular feedback—not just from having their most egregious sins named, but just as importantly, from getting some encouragement by seeing that not *everything* was unsatisfactory.<sup>3</sup>

### **Conclusion: Less of the Vertical, More of the Horizontal**

Everything I'm saying in this chapter can be seen in terms of a contrast between what I call a *vertical* and a *horizontal* emphasis in grading. We see the vertical emphasis in conventional A through F grades. (We see even steeper verticality when faculty members grade essays on a scale of 1 to 100—as they do in many law schools.) Conventional grades distinguish eleven levels of *pure quality*—quality that is entirely undefined and unarticulated: conventional grades constitute nothing but a vertical stack of levels—each one defined in

no other way than “better than the one below, worse than the one above”; it’s all numbers, no words; a yea/boo meter with eleven markings. This pure numerical verticality with no words or concepts is the source of the difficulty, disputes, and untrustworthiness in conventional grading. It is enormously hard to define or specify which essays should get an A and which ones an A- when we have no words or concepts or criteria. (“For an A, your paper can’t be just *really* good, it must be *really really* good.”)

I can summarize this chapter as two different movements away from pure verticality:

1. Minimal grading doesn’t add anything horizontal; it merely means *less* verticality: only three or so levels instead of eleven. That is, minimal grading doesn’t add any horizontal element—it is still nothing but a plain vertical line. But there is less relentless obsession with multiplicity of vertical distinctions. And defining the verticality with words or concepts (satisfactory, excellent) rather than just numbers gives a bit of relief from the verticality of pure numbers.

2. The use of criteria actually *adds* a horizontal dimension to grading. We are specifying two or more criteria at right angles to the vertical line of pure undefined quality. The fullest use of criteria involves making multiple judgments of quality. But even if we just specify criteria—naming them but not giving grades on them—we are still adding a significant horizontal dimension.

The choice between a vertical emphasis and a horizontal emphasis is stark. With the vertical emphasis, we are making a *single* difficult, sophisticated, evaluative decision along a *single* scale with multiple levels—but no words or definitions are involved. With a horizontal emphasis, we are making multiple decisions on multiple criteria—which are named—and the decisions are simpler, easier, and more believable. With the vertical emphasis, all the multiplicity is piled on top of itself—and undefined; with the horizontal emphasis, the multiplicity is laid out side by side—and defined.

We can also think of this as a choice between working with *quality* and *qualities*. To traffic in quality is to deal with a pure, unnamed, mystical essence. To traffic in qualities is to deal with admittedly crude and inexact entities—but at least readers can see what we are trying to evaluate.

My argument then is that if we use minimal grades, we are merely using less of the vertical dimension, but in doing so, we will make grading simpler for us, less questionable or dubious for students, and less disruptive of the teaching and learning climate. And if we use criteria and actually add some of the horizontal dimension, we will also make our grades much more meaningful, less magic or mystical. And we will be giving students valuable feedback on the strengths and weaknesses in their writing—feedback that they don’t get from conventional grades.

## Notes

1. Here are some activities that are often specified in contracts: attend class regularly; turn in assignments on time (to increase the benefit from discussions and lectures); revise cer-

tain papers; get rid of mistakes in spelling and grammar on final drafts; give written feedback on certain drafts to other members of the class; get written feedback from other members of the class (also perhaps people not in the class); turn in a *process log* with drafts and papers (containing elements such as a clear précis, a discussion of the writing process, an estimate of the strengths and weaknesses, and questions for the reader to answer in giving feedback).

It's probably most common to use a contract for a course grade of B—and then fall back on the question of excellent quality in deciding higher grades. But some teachers make contracts for higher grades containing additional contract requirements such as these: with each major paper, submit an outline that accurately shows the logic of the argument—the relation between main points, subsidiary points, and evidence; write a report on a book, lecture, movie, play not assigned for the course—showing how it sheds light on the course material; meet outside of class in small groups for certain tasks; make a presentation to the class or to some members—perhaps outside of class time; interview someone outside the class who has important expertise or experience in the realm we are studying—and make a report; give or get additional feedback on certain drafts; re-revise certain papers. Here are some more intriguing and speculative ways that teachers have tried to specify *activities* in writing to try to force students to write more productively: make sure that certain essays contain an element of exploration or perplexity or questioning—rather than just a summary of material or just an explanation of what is indisputable or obvious; certain essays must demonstrate the intersection of theoretical academic issues and some part of your personal experience—and how they shed light on each other.

2. A note about working out one's criteria: when students ask me, "What are you looking for?" I sometimes feel some annoyance (though I don't think my reaction is quite fair). But I enjoy it when I ask the question of myself: "What actually *are* the features in a piece of writing that make me value it?" If I try to answer this question in an insecure, normative way, I tie myself in knots: "What *ought* I to value in student essays?" But we are professionals in our fields, and so we get to ask the question in an *empirical* way: "What *do* I value in the writing in my field?" For there is no correct answer to the question, "What is good writing?"

This process of empirical self-examination can be intellectually fascinating. We learn to notice more clearly how we read—and this can even prompt some change in how we read. For example, some faculty members discover that they are judging on fewer criteria than they realized (for example, mostly on the basis of correct restatement of textbook and lecture material and correct mechanics)—and this realization leads them to attend to other criteria. Or they discover that they use different criteria for student writing than for professional writing (for example, in student writing they disapprove of the use of first-person writing or personal anecdote, but in published professional writing in their field they value it).

3. I need to add a note here about the vexed criterion of *mechanics* or *correctness* or *spelling and grammar*. Some teachers give it lots of weight—others not very much. Both positions are defensible. In working out my own view, I am helped again by invoking the *empirical* approach and asking myself, "How much do I value mechanics or correctness in the writing in my field?" The empirical principle helps me to see there is no single answer. That is, I don't value correctness very much in exploratory writing, drafts, informal reports, or e-mail—or at least not unless mistakes really impede my reading. I certainly don't care much about correctness in my own writing till I get to good drafts for readers I don't know. But I am put off when I notice mistakes in published writing, and so I work at correctness when I give good drafts to strangers or a wide audience.

I have derived my grading policy from this analysis. I don't use correctness as a criterion on drafts and informal writing, but I count it heavily on important papers that I treat as final drafts. (Usually, I set up assignments so there is a draft due well before the final due date—even if I don't read it. I tend to treat the draft as a kind of medium stakes assignment: it doesn't have to be done well—but the punishment is heavy if it is not done on time. This helps the high stakes paper a great deal.) For example, I might say that a high stakes final

draft is not eligible for the overall grade of excellent unless it has an excellent on correctness; and that a satisfactory is out of the question unless it is satisfactory on correctness.

Some students call this policy too stiff, and some colleagues say it isn't fair to students for whom English is a second language (ESL students). But it's workable once I make my reasoning clear to students: "I never penalize you for mistakes on drafts or on tests where you can't revise and get help. I'm not requiring you to know how to copyedit well—*on your own—without help*. But I *am* requiring you to learn to do whatever is necessary to get important papers well copyedited. Few of you can do it without help; *I* can't do it without help. Get help. What I am insisting on is an ability you will need for most teachers and most jobs: to do whatever is needed to turn in clean copy." It turns out that most ESL students are better at understanding the logic of this policy than most native speakers. I tell my students, "Here is another schizophrenia. On the one hand, I'm trying to teach you how to write on drafts without even thinking about correctness. On the other hand, I'm trying to teach you to worry a great deal about correctness in the last stages of preparing an important final draft." I'm not saying that this policy makes everything simple. I usually give students one "warning": the first time they turn in an important final draft that is badly copyedited, I usually just stop reading, reject it, count it as late, and say, "This isn't acceptable"—and let them fix it with only a small penalty. (They all get to have one late paper.) After that, I try to be as tough as I am pretending to be, for I believe they need to learn how to turn in clean copy.

## References

- Belanoff, P., and Dickson, M. *Portfolios: Process and Product*. Portsmouth, N.H.: Heinemann, Boynton/Cook, 1991.
- White, E. "Language and Reality in Writing Assessment." *College Composition and Communication*, 1990, 41, 187–200.
- Yancey, K. B. *Portfolios in the Writing Classroom: An Introduction*. Urbana, Ill.: National Council of Teachers of English, 1992.

*PETER ELBOW is professor of English and director of the writing program at the University of Massachusetts, Amherst.*

