

Do we need a single standard of value for institutional assessment? An essay response to Asao Inoue's "community-based assessment pedagogy"

Peter Elbow*

The Department of English, 170 Bartlett Hall, University of Massachusetts, Amherst, MA 01003, USA

Available online 7 November 2006

Abstract

Inoue [Inoue, A. B. (2005). Community-based assessment pedagogy. *Assessing Writing: An International Journal*, 3, 208–238] sets up a radically experimental writing class as a kind of laboratory of assessment. He seeks to avoid the standard situation where a teacher unilaterally assesses and grades student writing, using only his or her own criteria or standard. His premise is that value in writing is socially constructed, and so he gets students to enact a social, communal process to work out the criteria for effective writing. In this way he gets students to take full responsibility for assessing and even grading each other.

But when he requires his students to agree on a *single* model of good writing (a "rubric" with various dimensions), I think he misunderstands how value is actually socially constructed. In fact we live in a world with various models of good or effective writing. Given, however, that he is working in an institutional setting, his approach seems readily understandable. How can feedback be coherent or final grades be fair if they are based on *multiple* and *competing* models of goodness? This sounds like a rhetorical question, but in fact I use this essay to argue that it is not only feasible but desirable to use multiple models of value in an institutional setting.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Writing assessment; Grading criteria; College writing; Social construction; Value

In a remarkably valuable essay, Inoue (2005) describes a radically experimental college writing class. He designed it not just to teach writing but as a kind of laboratory to explore the assessing and grading of writing. He takes students out from under the pressure of teacher assessment and grading (what he calls a "hegemonic discourse") and put assessment entirely in their hands for

* Tel.: +1 413 545 4339; fax: +1 413 545 3880.

E-mail address: elbow@english.umass.edu.

what he calls “community-based assessment.” “Our class writing isn’t about what I want — it’s about what the class can agree they want. . .” (p. 221).

Over the 15 weeks, he gets his students to develop – gradually and carefully – a set of criteria for assessing and even grading each others’ essays (he calls these criteria “the rubric”). He insists that they use these communally derived criteria to assess and grade each others’ papers. In this way he undercuts the pervasive and depressing tendency for students to feel dependent on teachers and to withdraw their own commitment to hard thinking about the nature of excellence in writing. Conventional grading seduces too many students into thoughtlessness. Sometimes it is knee jerk acceptance of our grades and feedback. (“Of course my teacher knows best. Her good advice will improve my writing and my grade.”) Sometimes it is a facile, pragmatic skepticism. (“What my teacher wants is just academic [or intellectual or abstract or politically correct] bullshit. But of course I’ll go along with her — I don’t want to hurt my grade.”)

Inoue’s class is adventuresome and revolutionary, but it is not just a brave journey into anarchy. He imposes carefully planned processes and activities that get students to take on their assessment responsibilities in an impressively thoughtful way. Students must read each others’ writing, and published writing too, with careful attention so as to think hard about criteria for successful writing. They must work out agreements that are continually negotiated.

On the one hand, he sets up *inductive* situations where students bring in examples of writing they admire, and from these they try to derive and articulate what makes them effective as writing. But then on the other hand, he builds in a *deductive* process where the rubber hits the road: students must apply these criteria to each others’ essays and produce feedback and even grades. Through this process, he gets students to think hard about criteria for good writing: what are good standards and how can we apply them? He wants a course where students “articulate and theorize assessment practices” (p. 234). As an important side-benefit, his course puts students in a better position to understand and deal with the (often unarticulated) criteria they will encounter at the hands of other teachers.¹

1. Where does value come from?

Behind the common question, “What is good writing?” there lurks the larger question of how value is created or bestowed in the world (or in various communities in the world). This question is too often evaded, but Inoue’s title phrase, “Community-Based Assessment,” signals that his experiment in institutional grading and assessing is also a laboratory to explore this larger

¹ I’m struck at how his class is a story of skillful management. Inoue keeps a remarkable number of balls in play. I could teach a good course about Chaucer with no management skills at all, but I could not even make it through half a semester of a writing course without lots of management skill that I have had to struggle hard to learn. Inoue shows himself a master. (His students are mostly third year students. I am interested in how his approach could be applied to the ubiquitous first year writing course.) His skillful management gives the lie to the widespread assumption that there is a binary opposition between ‘student centered’ and ‘teacher centered’. A good ‘student centered class’ often takes a lot of teacher authority. (“It takes a lot of resources to keep Gandhi in poverty,” remarked someone once in his entourage.) Teachers do not need much authority for ‘teacher centered’ classrooms where they just lecture or run a standard discussion and the default pedagogical flywheel turns of its own inertia. Inoue’s class is student centered. The teacher needs to play a strong and central role in order to get students to do what they are not accustomed to doing — in this case taking over the process of figuring out standards and then applying them to each others’ papers in a rational and well organized fashion. I admire how Inoue applies authority in an open, non-defensive but deft way. Students often resist workshop activities where they must depend on each other and not get help from the teacher; they often resist negotiating agreements with classmates who have different values or styles.

question. I write to praise how far he comes in his remarkable journey of pedagogical and theoretic adventure. But I also write in order to explore how he (and those of us following his lead) can go farther toward his ambitious goals.

I have taken the bait that Inoue's essay provides and I have tried to think more about the sources of value. I will explore three dimensions. For each dimension, I will try to show correlations between what happens in the world and his illustrative class or laboratory or model.

1.1. Value bestowed by individuals and value bestowed by groups

When individuals call something good or bad we can often see how their values mimic the values of the group they are in. But sometimes individuals seem to express a more idiosyncratic value — whether we explain this as genuine self-authorized autonomy or merely some unique mixture of influences from various groups. I highlight this distinction between individual and group because I want to explore the fruitful dialectic in Inoue's class between individual value judgments and group votes.

1.1.1. Individuals

When individual teachers give grades, they 'endow value' in certain kinds of writing and not in others. Teachers — and other individuals like parents, friends, or critics — can get us to decide there is value or excellence where we did not see it before. An older brother or friend often gets a young man to see more value in beer and less in soda or lemonade. T.S. Eliot got many readers to see more value in Donne and less value in Tennyson. When editors of journals decide which articles to publish, they endow certain kinds of writing with value. Even when they use a system of blind reviewing, editors often express their values in various ways, for example by deciding which 'blind' reviewers to use for a given submission.

Smith (1992), drawing on his work with colleagues at the University of Pittsburgh, changed my sense of value not in writing itself but in the process of grading placement tests. He convinced me that raters should not be trying so much for an "accurate score" for the text itself; rather they should be looking *through* the text to try to see the writing abilities of the writer behind it. The judgment he asks of readers is harder and more questionable; it's far easier and more accurate to score a text. But the quality of the text is not what we most need to know for placement. We need to know the skills and abilities of writer who produced it. (Here is an instance of that perennial problem: do we measure what is easier to measure or try to measure what we actually need to know?) And Smith is surely right that this tricky difficult judgment is best made by teachers of the courses into which students will be placed. These teachers need to read papers with this simple question in mind: "Do I think this student belongs in my course?" not "What is the correct score for this paper?"

1.1.2. Groups

Agencies like the Educational Testing Service use their tremendous authority to give value to certain kinds of writing. The Association of Writing Program Administrators, with their collaboratively produced "Outcomes Statement," bestowed value on certain outcomes for the teaching of first year writing. Professional organizations, journals and other publications do the same thing. *The New Yorker* gave value in the 1950s and 1960s to so-called "*New Yorker stories*." At first, many readers did not feel these stories were good. "*Nothing happens. And there is no ending.*" Yet through the power of *The New Yorker*, many readers came to bestow value (had their eyes opened to value?) where they had not seen it before.

The larger culture or ‘the tradition’ can function as a ‘group’: people like to point to how, over centuries, ‘the culture’ seems to have agreed that Shakespeare produced ‘good writing’ – even while it seemed unable to make up its mind about so many other writers whose stock rises and falls over time. There was a group of women poets in the 19th century who were more popular and considered better than Wordsworth, Keats, etc, yet somehow their names disappeared from the canon of English literature. Frost and Raymond Chandler are now highly valued but used to be dismissed. It is always a story of power being deployed (whether or not we think it is a story about inherent value or socially constructed value).²

I admire how Inoue creates a productive dialectic between the individual and the group in the process of figuring out how to bestow value on each others’ writing. First come the judgments of individual students as they formulate criteria. And throughout the semester, students continue to give individual comments and responses to each others’ writing as they try to apply their own criteria to texts. But Inoue sets up a carefully orchestrated structure whereby all the individuals negotiate and vote so that the final version of the criteria for good writing is a *corporate document*. As such it exerts enormous corporate authority back upon individuals in the class. Individuals must abide by it, even when they do not like it.

1.2. *Value bestowed by conscious analysis and value bestowed by unexamined reactions*

1.2.1. *Conscious analysis*

Teachers often meet together to deploy conscious analysis for figuring out what makes writing good — at least for their purposes. The same thing can go on in a ‘norming session’ for exam readers. The history of literary criticism (like art and music criticism) could be seen as the story of smart people trying to analyze works as thoughtfully and consciously as they can in order to work out explicit criteria of excellence or value.

I have a vivid memory of my own conscious analysis of value in college. I heard the Budapest String Quartet play a Bartok quartet and found it to be mere chaos – unpleasant at that. I loved music and was taking a music survey course at the time and had to write a term paper on a work not covered in class. I chose that Bartok quartet. I was influenced by the fact that some people I admired praised Bartok – and the Budapest Quartet had enormous authority in an era where there were few eminent quartets in the U.S. For the paper I engaged in conscious analysis. I looked for themes and structures and so on, and was impressed to discover that the piece was not the chaos I had assumed from my non-analytic first listening.

1.2.2. *Unexamined reactions*

Unexamined, non-analytic reactions often determine which books and songs and paintings become popular (though not always). But we ‘intellectuals’ are not immune from the influence of unexamined reactions or even unexamined prejudices. Like many readers of this essay, I have a visceral prejudice against five-paragraph-essays. Periodically I try to remind myself that the form does not in itself preclude good thinking and even aesthetic pleasure. I also have a knee-jerk

² Hirsch (1977) made a serious stab at trying for an objective measure of goodness in transactional prose. He thought he could calculate effectiveness by simply measuring how long it took for readers to understand different texts. But the nice idea refused to yield good numbers: the same text was quick for some and slow for others. He ran afoul of the deepest problem in figuring out goodness: different readers bring different mentalities to the text. It was this experience that led him to his obsession with “cultural literacy” — a massive project of trying to get *all* readers to have enough of the right stuff to bring to what they read.

prejudice against essays that start off with a definition from the dictionary. For a telling indictment of the thoughtless prejudice that many academics have against the sentimental dimension in writing, especially in student writing, see [Newkirk \(1997\)](#).

I admire how Shakespeare and Chaucer are ‘high’ and ‘great’ writers and yet frankly try in much of their work to please the unexamined non-analytic reactions of a broad audience. Indeed unexamined reactions can be acute – sometimes decidedly more fine-grained than conscious thinking. Our unconscious knowledge of grammar is far more extensive and acute than our conscious knowledge of it.

1.2.3. An integrative process that combines the conscious and the unexamined

What I remember most about my work on that Bartok quartet was an interestingly mixed process. That is, as I tried to write my paper, I did not just look at the score to find patterns; in the process I also simply listened to the work over and over on a phonograph record. I engaged in a kind of ‘believing game’ process of simply trying to *hear* what it is that smart people heard that I could not hear. Gradually I heard the sounds completely differently; it was as though I heard new sounds, certainly new structures. Instead of *noise*, I experienced beauty on the most visceral personal level.

In his course, Inoue naturally pushes his students hard to engage in conscious analysis as they slowly, deliberately, and collaboratively work out their criteria for good writing. He does not want the assessment process to be determined by mere gut reactions or by the personality or popularity of certain students or published writers. He’s fighting against the path of least resistance for students (and the rest of us): unexamined reactions. But on the other hand, he also makes an important space for gut reactions, recognizing how they often serve as excellent ingredients for conscious analysis of value – sometimes showing us what conscious analysis overlooked.

1.3. Value bestowed by induction and value bestowed by deduction

1.3.1. Induction

In some testing situations, the administrators derive their criteria inductively from the reactions and discussion of readers. Good critics often start from an inductive examination of reactions. [Gardner \(1956\)](#) proclaimed a striking principle I have always wanted to believe but continue to ponder: that no critical analysis – however smart and impressive it is – should be trusted if it violates the experience of most readers or viewers.

1.3.2. Deduction

Deduction is a matter of holding up criteria or principles, whether long-established or new, to see if a given text fits them. What *ought* good writing to look like? According to some theorists and many teachers, a good story must have a ‘crisis’ and a ‘resolution’. All good essays should open with an overview of the whole, and they should never leave readers feeling lost.

The deductive principles that I just enumerated for stories or essays illustrate the dangers in deductive calculations of value. Those principles rule out many stories and essays that many good readers find undeniably excellent. The deductive process is a particular danger for teachers. We often read essays or stories in stacks of 25—and our purpose in reading them is often to judge or grade rather than to learn or enjoy. This way of reading tends to drag us into a process of holding up texts against a template of criteria in our mind and merely seeing how the essays fit. Only teachers, exam readers, and contest judges read parallel texts in stacks of 25. What pleases them

as they read in this fashion is often different from what pleases readers who read texts one at a time to learn or enjoy.

Aristotle was being inductive in his *Poetics*. He wanted to see what characterized the plays that won the contests that were decided by the votes of Athenian citizens – votes often derived from unexamined reactions. But his inductive findings (e.g., that good plots have a crisis or turning – or that the action takes place off stage – and in twenty-four hours) have been notoriously used as deductive requirements through the ages (especially strongly in the 18th century). He became a single figure with enormous authority to shape criteria of value.

The common misuse of Aristotelian hypotheses represents what we might call a perverse, *one-way* link between initial induction and rigid deduction. What is needed is a productive, *two-way* dialectic between induction and deduction. Each can serve as a corrective to the other. That is, our inductive reactions can be corrected by paying attention to criteria: “*I hated that essay. But now, as I look at various criteria, I see that the problem was a pervasive and annoying voice. My inductive reaction blinded me to other dimensions or criteria where the essay was strong.*” By the same token, our deductive criteria can be corrected by paying attention to inductive reactions: “*That essay clearly fails according to our criteria. Yet I experience it as terrific. You do too. Hmm. Back to the drawing boards for our criteria.*”

2. A confusion about how valuing works in the world

In effect, Inoue is insisting that the concept of good writing is socially constructed. Thus, he stresses the terms “community-based” and “communal”, stressing “writing as a set of conventions” (p. 220) and insisting that “writing conventions are contextual and communally developed; that they often evolve over time. . .” (p. 232). He builds an elaborate process of social construction for figuring out values: “The rubric is central because it forms both our discussions of writing as a set of conventions, and it is used to assess the writing of the class” (p. 220).

Surely he is right about this general principle. What could be more socially constructed than ‘goodness in writing’? The very concept of ‘good writing’ in most minds *means* writing that *people* call good. Arguments for an external and eternal objective Platonic standard for ‘good writing’ are rare.³

But despite Inoue’s brilliant design and execution and the obviously deep learning we see students taking from his class, I have a serious reservation about how he designs his laboratory for studying value. The problem comes from a central premise that underlies the whole operation: he insists on a *single* unified set of criteria for judging all writing in the course – a single picture of effective writing. Even final course grades derive from this corporate agreement. When he makes the class vote and agree on a single standard of effective writing and stick to it for all

³ I should note that Inoue does not so much ask students to define *excellence* in writing as *effectiveness* — what works. When people concentrate on ‘true excellence’ in an exalted sense — pure beauty or pure truth — they are often more tempted to talk about an eternal standard. But when it comes to what works, we are surely in the realm where readers decide. Of course not everything is socially constructed — or at least to the same degree. At one end of the continuum, ‘goodness in writing’, is way out there where things are highly (wholly?) socially constructed. In contrast, ‘goodness in mathematics’ depends importantly on how many apples you end up with when you put two apples with two other apples — and depends also on certain structures of impersonal logic. Rocks and stones and trees have a socially constructed *dimension* that stems from how we conceive of them — but also a not-socially-constructed dimension in what persists across widely different cultures and ways of conceiving them. Thus debates over social construction are unhelpful when they take the form of a yes/no argument about whether ‘everything’ is socially constructed or not. It’s far more accurate, interesting, and productive to examine the *degree to which* and the *way in which* different things are socially constructed.

evaluation, he seems to be saying, “This is a picture of how valuing works in our culture or our institutions. You may not like it, but that is how power is deployed: it flows from group agreements.”

But the social construction of value is more complicated than he implies in his laboratory. There are crucial differences between how, on the one hand, *meanings* and *conventions* are communally created – versus, on the other hand, how *value* is created. Yes, meaning in language is communally agreed upon: natural language depends on subtle, largely unspoken, communally negotiated agreement about the main denotative meaning(s) of each word (despite different connotations that every word has for different speakers). And these agreements always involve a dimension of power (in a process I described in 1973 on pages 151–157 of *Writing Without Teachers*). *Conventions* too, represent social agreements, and conventions certainly play a role in what is called acceptable writing – whether it is a mechanical matter like the use of commas and capitalization, or a subtler matter like what is the right structure or what counts as an argument in certain genres. Conventions get their name because they are “conventional” – things that people agree on. But when it comes to the social construction of *value* – what is good writing – the story is more complicated. The *construction* job that people engage in when they create value is a hard hat job. “Construction workers” cannot agree and are often hit over the head.

And yet Inoue shows that he knows all about conflict. His class is a perfect laboratory where values collide. It is as though he issues hard hats. He shows this most vividly and touchingly in the story of Elizabeth. She reads a paper by a classmate (late in the course, it seems) and finds it terrific. But unfortunately, the corporate criteria of good writing disqualify it from being terrific. Inoue quotes her: “It is difficult to grade a great paper that just did not fit our criteria” (p. 228).

Here she hits the empirical nail on the head; she locates exactly the place where the dialectic between induction and deduction ought to be re-opened. Yes, she may be crazy, she may love that paper for terrible reasons, but Inoue’s course process does not allow people even to explore the merits of her response. There is no place (at this point in the course) to explore whether perhaps that paper reveals a weakness in the corporate criteria. Perhaps the paper has some dimension of goodness that they had not figured out or had mistakenly dismissed – a feature of goodness they might even be talked into if the door had been opened.

The theoretical problem is bluntly revealed by Inoue’s own words at this point: “she frames the ‘badness’ of the paper . . . not in terms of ‘badness’ but as strengths that just do not ‘fit our criteria’” (p. 228). What is remarkable here is the word “badness.” It is not Elizabeth’s word – she loved the paper – it is Inoue’s word. He is so committed to a single corporate set of criteria that must be applied deductively that he feels entitled to use the word “bad” (in scare quotes) for her experience of goodness. (Possibly the word “bad” was used by Elizabeth in some conversation, but this would not affect my point: it would simply be a case of her feeling obliged to conclude that what she experienced as good was in fact bad.)

In fact, Inoue *celebrates* this situation where a single value must be wrested from conflicting values. He celebrates individual and dissenting voices:

When reflected upon and discussed publicly, writing and meaning making become less about what is good for the individual and more about what’s good for the class. Yet these same reflective activities can also preserve individual resistance to discourse conventions set up by the class (the rubrics) by allowing spaces for dissonant voices and uncensored questions. In fact, this happens often in my classes, and I encourage it because it gives us a chance to talk about why we believe what we believe, about hegemony, the power of language and its *conventions* over us and our ways of knowing. It connects the self-as-writer

to the self-as-stakeholder-in-community. It's an attention to the "us" yet not forgetting the "me" (p. 235 my emphasis)

(Again he invokes the fact of *convention* to defend using a single model for good writing.)

My main argument here is this: the process of wresting agreement from disagreement is a deeply incomplete picture of how value works in the world. Admittedly, there are important agreements within certain subgroups of society or culture. ETS and other testing agencies, for example, work out corporate agreements about what makes a '5' essay; graders of a local placement exam often agree on what makes a 'basic writing' paper. But these agreements often differ from one institution to another – despite some large overlaps. And these very agreements testify to the failure of any *unified* social construction, since the testers have to work so hard to 'calibrate' readers – which means forcing them to check their own values at the door. Broad (2003) made a rich and extensive study of the actual evaluative talk by teachers engaged in grading portfolios. His analysis documents extensively how humans – even though they are all teachers in the same first year program – differ in their reactions and their standards.⁴

So even though teachers and testers and others can work out socially constructed agreements in particular subgroups of society, these agreements belie the larger picture of how humans bestow value and how value happens in the world – the picture of where the concept of 'good writing' comes from. Our concept of good writing is socially constructed – but *only* if we understand that there are multiple and competing social constructions of good writing – even in the same community.

There is a more technical and authoritative way of making this point. Interestingly enough, it is made definitively in an essay printed right next to Inoue's in the same issue of *Assessing Writing*. Slomp and Fuite (2005) analyze in detail the relationship between reliability and validity in testing. They show how there is an inherently inverse relationship between reliability and validity. As reliability improves, validity degrades – and vice versa. (This was a claim I made in my "Foreword" [Elbow, 1991], especially pp. xii–xiii.) Slomp and Fuite make their case with great care, using both empirical research and a theoretical/mathematical argument. And the inverse relationship between validity and reliability they show is not simple: the degree of interference between validity and reliability *varies* depending on what is being tested (the 'construct'). If the construct is very simple – such as a test only of capitalization – they show how scores can be remarkably high for *both* reliability and validity. But in a test of anything as complex as 'good writing', the trade off between validity and reliability is severe. Their research is complex, but they can sum up the bottom line in simple common sense terms: "the construct 'good writing' has yet to be fully defined and agreed upon" (p. 204). Herrnstein Smith (1988) has written what seems to me the definitive theoretical and philosophical study of the 'contingency of value'.

In short, disagreement is common when it comes to any account of good writing that is precise enough for assessment purposes – much less grading – even within small groups or communities.

⁴ In a response to a draft of this paper, Ed White argues that we see something like single standards for good writing if we control better for genre, purpose, context. For example, "school writing, like most writing in my world, is on a specific topic with a defined purpose for a particular audience" (personal communication, 22 March 2006). I agree that if the purpose is sufficiently narrow, there will be much more agreement as to success, for example a school assignment to summarize a published essay, or a workplace memo to summarize a meeting. Yet even on these narrow tasks, we sometimes see disagreement about whether a summary was accurate. And most writing tasks involve larger purposes, for example to persuade or to analyze, and disagreement is widespread about success meeting those goals.

When Inoue speaks of “assessing writing in a real and tangible community” (p. 230) he implies that communities produce agreement about value – when in fact they produce agreement only about conventions.

My claim that there is no agreement about good writing will sound anarchic or individualistic to readers who are too beguiled by loose talk about speech communities and the social construction of conventions. Many people say things like, “We in our speech community or discipline agree about what good writing is.” Teachers of writing like to say to students, “The academic community agrees what good (academic) writing is.” But those agreements do not really carry much weight for individual members – and they do not cut very fine. If we use the right lens, we can stress the genuine commonalities or agreements in valuing that exist in disciplines, but this lens shatters every time students submit the same paper to different members even of the same department and get different grades. Disagreement is more profound when you cut across departments and move outside the academy. (See Bizzell, 2002 for a case where the premier journal in history had enormous disagreement and struggle about whether to publish an essay by one of the leading scholars in the discipline. We see agreements when Nobel judges, and other committees, award prizes but these are temporary enforced pragmatic compromises. And we know that judges often hang on to their individual disagreements – sometimes even airing them in the press afterwards.)⁵

3. Why does inoue insist on a single standard?

Inoue is highly sophisticated and well read in his explorations of evaluation. So why is he unrealistic about this one dimension of his otherwise realistic class? It is not for lack of courage to tread where teachers usually fear to go: removing himself entirely from any official evaluation of his students’ writing. His courage did not impede him; it helped him be remarkably clear sighted and accurate about most of this course. I have two hypotheses for why he insisted on a single standard.

First, I suspect Inoue wanted his students to know what it feels like trying to grade fairly – to have (in his phrase) the “hegemonic power” and to feel what it is like to wield standards that one has generated or has power over. And yet he cannot let students generate official grades unless they agree on one standard: it would seem too anarchic in an institutional setting if, for example, each student had the teacher’s experience of solitarily giving a final grade based on his or her own standards (for example each student could grade the one who comes next in the alphabet). He cannot bring himself to give his students this universal teacher experience because he sees too many problems in it. He feels obliged to give his students an experience that is far too rational to let them experience what teachers feel. His students’ evaluations are grounded in carefully derived and publicly negotiated collaborative unanimity. No doubt Inoue helps his students see this process as a useful *critique* of the way teacher grades and exam scores are usually arrived at.

⁵ When Inoue insists on a single model or rubric for good writing, this also leads to an unrealistic limit on the topics students can choose to write about. For as the class developed their rubric (entirely without Inoue pulling any strings behind the curtain, as so many teachers might be tempted to do), they understandably conclude that good writing should be “interesting.” As the class applies this criterion they end up down-ranking pieces that don’t interest many members. (Maybe the problem is not with the criterion but with how they apply it. I don’t think “interesting” *has* to mean “interesting only to the majority”? Writers in the world often write to a tiny audience and produce writing we have no trouble calling good – sometimes even “interesting.”)

My second hypothesis is more theoretical and psychological – and it is where I want to put my emphasis. When I wrote *Writing Without Teachers*, my shrewdest critic told me, “This still smells of teachers.” Inoue’s critique of institutional assessment still smells of institutions. He cannot (as I could not) extricate himself from the context he was working in. Because he is been breathing it for so long, he is still stuck with a central assumption underlying virtually all institutional evaluation and assessment: the need for a single verdict as to quality. After all, he immersed himself deeply in the literature of assessment. And in addition he designed his course while an advanced graduate student – which puts him doubly in the belly of the beast: he is giving institutional grades and *being* graded at the same time.

It is my hypothesis, then, that this pervasive assumption of the need for a single standard interacted with his accurate idea that value is socially constructed: those two assumptions fermented together to yield the idea that the class has to agree on a single standard or rubric. I care about this issue because I believe that these two assumptions ferment throughout the composition community – indeed throughout the whole assessment community – to produce the same sour result: a confusion between ‘value’ and ‘a single verdict’. Discussions of evaluation, assessment, and grading are too seldom built out of an understanding of the essential complexity of actual value in the world: where it comes from and how it functions. In the world – even in many small local pockets of the world – there is no agreement about good writing. And yet the story of value in the world is not *merely* the play of arbitrary individualisms.

It is the same truth that Slomp and Fuite found in their sophisticated study of the relationship between validity and reality. Until we are clear on this point, we will mislead people about the social construction of value. Yes, ‘good writing’ is a socially constructed concept. *But* (a) the complex construction process yields a wide variety of simultaneous results. *But* (again) (b) this does not mean we cannot talk rationally and theorize about value. That is, we do not have to throw up our hands and fall in line with defensive freshmen and cynical intellectuals who say, “That is your value, this is my value, there is no use talking about value, value has no meaning.” In fact it is possible to engage in productive thinking and theorizing about value and about where people agree and where they do not.

4. Let us push farther down the brave and helpful path that Inoue forged

I hear in my head the following plausible response to my case so far:

You’re just stating the obvious. Of course the larger world of writing is different from the institutional world of teaching. Of course there is no single standard for good writing in the world. But in an institution, we cannot have fairness unless a teacher uses a single standard for grading all her students; and institutional exams cannot be fair unless they are graded with a single standard; and when a bunch of teachers teach the same course, they should at least try to use a similar standard.

Plausible yes. But I disagree. I will end my essay by imagining how *even in institutions*, we can give students an assessment experience that is both fairer and more “liberatory” (Inoue’s word – he also speaks of “education as the practice of freedom” rather than one to “reinforce domination” [p. 237]). And in this process, we can give them a better conceptual understanding of the nature of assessment and valuing. I am making two concrete suggestions here: how we could adjust Inoue’s process by which students negotiate value in writing; and how we could determine fair course grades in a way that fits the principles I am arguing for.

5. A social construction of value that does not enforce unanimity

Only a small change is needed in the negotiating process. We would have students use exactly the powerful and well planned processes of social, communal induction and deduction that Inoue sets up. We would simply refrain from forcing them to agree on *one* set of criteria or one rubric. As a result, they would probably end up with a small or large handful of sets of criteria or rubrics: perhaps quite a few students would agree on one standard; then a few smaller groups would agree on different standards; and there might be a number of loners who do not agree with anyone else. (Of course it is possible, some semesters that the whole class might come to agreement in the absence of any rule that forces them to do so.)

In effect, this process would help them, as a community, create a more valid laboratory or model for the world of writing. The world (including of course various academic sections of the world) is rife with conflicting criteria of good writing. And yet the world and the academy are full of people thinking carefully about value and comparing notes with others. We will give our students a precious gift by helping them see and enact this complex process.

Admittedly most students already see the divergent standards among teachers – even in the same department – but too often they see it only through a cynical lens. “*All teachers have their own personal prejudices. ‘Good writing’ is just a crap shoot concept.*” Such a lens devalues serious thinking about the notion of excellence as a complex social construction. And if we want to help students learn to read and write better, we need to help them see competing standards as a *positive resource* for understanding quality – and a testimony to the complexity and diversity both in pieces of writing and in pieces of humanity. Multiple lenses for excellence permit not just richer, but also *more accurate* perceptions of texts. A single lens always hides or distorts aspects of what is being looked at.

So I am not proposing a process that is less communal. Students would continue to work just as hard to figure out a definition of good writing. But they would not have to agree with each other. It would be just as communal because (like Inoue) I would use my authority as teacher to stack the process to build community. Just as with Inoue’s process, each student would have to get responses from other students about where his or her definition makes sense and does not.

In this careful process, then, students have to work in groups and in fact be encouraged to find others with similar criteria for ‘good writing’ and work within these local groups of like-minded readers. A communal process stripped of the requirement for agreement, would nevertheless help the class to see *potential* agreements – *unforced* agreements in their thinking – while helping them articulate where they disagree. This would help them sharpen and clarify criteria.

Someone like Elizabeth who is fighting for a minority definition of good writing would be obliged to listen to everyone else’s disagreement with her and hear their conflicting definitions of good writing. But if she held fast to her criteria after this process of sharing and listening and discussing, she would get to have her own standard. The social process might even lead a couple of others to join her. But whether a definition of good writing represents lots of students or only one, it would have to be worked through a communal process of development, feedback, trial application to texts, further development or revision, further feedback, and further application to texts. Thus the process would be just like Inoue’s in its social foundation and the valuable dialectic alternation between induction and deduction.

The benefits of this change would be both theoretical and practical. Theoretical benefits: With only this small but powerful adjustment, students would get a more sophisticated and accurate understanding of the very concept of good writing and of how valuing actually works in the world. This would be particularly helpful in preparing them to write for other teachers and to see

those teachers' evaluative judgments in the most useful perspective. If students are led through this process, they will discover not only that people differ in what they value, but that many or perhaps even most of these conflicting values make a genuine kind of sense – a sense they had been blind to. Good writing is such a complex construct that there are myriad kinds of 'goodness'. Students would see (as Elizabeth incipiently sees) that different criteria – different definitions of good writing – highlight different features of any given text and help us see more about it.

Practical benefits: Students would learn to give and get more useful feedback on their writing. If we want to be good readers – and good writers! – we need to be better at using multiple lenses – at seeing and appreciating a variety of competing criteria. If we read only through one lens that focuses only on one set of criteria for goodness, we blind ourselves to features of a text that affect some or even many readers. We blind ourselves to excellences and faults that need to be recognized – not only for valid judgments but also for our own pleasure and edification. For example, some criteria work better for certain texts than for others. Certain criteria bring out excellences we did not notice. For example, an *ethos*-oriented reading ('personal' or 'voice-oriented') will show strengths in some texts and weaknesses in others – just as a *logos*-oriented (logic-driven) reading will show strengths in some texts and weaknesses in others.

This metaphor of lens (what Burke calls a "terministic screen") leads interestingly to two opposite practical principles. On the one hand, we should always try to find the 'best lens' – the lens that shows most of what is there. Thus, if we want to understand and do justice to how a work is organized, we should use the organization lens that best fits it. A five-paragraph essay organization lens blinds us to subtler organizations that might actually be functioning in an essay.

But I have always been interested – in addition, not instead – in the opposite principle: the value of 'wrong lenses'. It usually takes a wrong lens to find trace elements in a text that most readers or viewers do not notice but which can increase our understanding and pleasure. And not just that: trace elements that most people cannot see when they use an 'appropriate lens' sometimes influence how even those readers react to the text. The commonest form this takes is when readers do not like a text and give all kinds of reasons, but are actually turned off by certain trace elements (for example a bit of condescension, sexism, self-pity). My interest in wrong lenses leads to my interest in the 'believing game'.

If this process of using multiple lenses sounds too anarchic, it is worth realizing that it is actually business as usual in most of our classrooms. That is, whenever we set up groups for peer response, students give each other evaluative feedback based on their own divergent standards of good writing. What is different, of course, is that Inoue has ambitiously insisted that students work to make their criteria *conscious* and *explicit* – in a slow motion process that gets them to interrogate what they formerly just took for granted.⁶

⁶ When in 1973 I suggested using peer response (to a composition world that mostly did not use it), I was nervous about the way it involved multiple and divergent lenses. My long appendix essay about the believing game was an extended attempt to justify what might be seen as intellectual chaos. That essay (and succeeding essays I have written exploring the notion further) are attempts to defend the value of multiple lenses on epistemological grounds. A practical application. When I was director of the writing program at SUNY Stony Brook and conducted my first norming sessions for portfolio readings, I saw how much useless energy and adversarial feelings we generated when we tried to insist that readers agree in their judgments. I worked out a corrective process based on the believing and doubting games. As we looked at a sample borderline portfolio, first I pushed all the readers to play the believing game: "Everyone, please try as hard as you can to find as many reasons as possible why this portfolio should pass — no matter how you actually feel about it." This yielded lots of strong and interesting reasons for passing — some that readers had not noticed at first. (Playful, far fetched and even satiric arguments were often quite fruitful). But then: "Now, please, everyone try to find as many reasons as possible why this portfolio should fail." Again, lots of good perceptions and arguments. In short, I set up a communal process to

6. But how can we get fair course grades without a single standard for good writing?

I hope I have demonstrated, then, that multiple criteria are no problem when it comes to giving students feedback and evaluation of their writing – even in an institutional setting. When students see their writing through the lenses of various readers and various conceptions of good writing, they can see their own texts better and make better strategic calculations about how to revise for a world of disparate readers.

But what about course grades? How can they be fair without a single standard of excellence? I will argue here, in closing, that in fact there are other ways to get fair course grades besides basing them on attempts to measure the quality of students' writing against a single standard. But my argument rests on a negative premise: using a single standard cannot give us fair course grades. Thus the goal for grading cannot be perfect fairness. My goal is to get course grades that are at least as fair as conventional course grades — i.e., no less unfair. To support this negative premise, I list five serious impediments that undermine the fairness of conventional course grades based on a single standard of writing quality:

- The first impediment is the main argument of this paper: any single conception of 'good writing' is an inadequate representation of quality – a faulty window on the multifarious picture of how humans value writing. As Fuite and Slomp say, "the construct 'good writing' has yet to be fully defined and agreed upon" (p. 204) – often even between teachers of the same course. Thus, course grades based on a single conception of 'good writing' are invalid reflections of the construct. Note that the widely accepted arguments against IQ tests come into play here. We do not make intelligence tests fairer by deciding on one standard or construct. It may be a big problem that there are multiple standards for intelligence, but fairness means recognizing them, not sweeping them under the rug. (Yes, we might want to test people on just one dimension, for example memory or creativity, but then we cannot call this 'intelligence'. So with writing: grades are fairer if we measure only mechanics or organization, but that is not 'writing'.)
- Conventional course grades like A through F are one-dimensional signs or symbols. A one dimensional number will always lack both reliability and validity when it attempts to represent the quality of multidimensional performances. That is, a single number cannot fairly represent the *combined* values of the various dimensions in a piece of writing or a portfolio (e.g., ideas, organization, sentence clarity, surface features, etc.). Different graders will come up with different one-dimensional grades as they give different weightings to different dimensions of the writing. In addition, different *readers* of a transcript will infer different meanings when they see B. (Inoue shrewdly puts a heavy emphasis on a multidimensional "rubric" in his evaluation process—though of course in the end, he has to reduce the verdicts to the one-dimensional grades required by the registrar.)

harness a community of teachers to learn how to see a portfolio better by seeing it through two lenses — as passing and as failing. This helped us see more of what was in the text. Only after that process did I ask everyone to vote their final verdict: pass or fail. But this vote was not binding — whether it showed a clear majority or an even split. The crucial point, both practically and theoretically, is that it is not worth trying to badger people into agreement. The final vote gave everyone a better picture of "good writing" in our community — the mixture of values that constitutes our community. And the picture is more accurate than with the normal authoritarian/competitive situation when people fight to make their lens "win." "Outliers" with idiosyncratic standards of good writing will see that they are outliers and most of them will feel at least a bit of pressure not to make their own portfolio judgments be *too* different from those of their colleagues. Forcing people to "bow their knee to the majority" does not really do much to change their idiosyncratic thinking. About the believing and doubting games, see my Elbow (1973), Elbow (1986), and Elbow (2005).

- Most course grades are unilateral judgments by one person with no second opinion. Of course there is nothing wrong with this in itself. The problem comes when that single judgment is used as the *only* judgment and as an official grade, as though it has more inherent validity than other judgments based on other standards. Inoue may insist on a single standard of excellence, but at least he does not let that standard be the invention of just one person.
- Few teacher grades are derived from a slow careful process like the one that Inoue sets in place. His students communally devise and publicly articulate a multi-dimensional rubric in a process that involves a dialectic between induction and deduction.
- Finally, even if none of those four impediments existed – even if a teacher *could* magically calculate grades that are *perfectly fair* representations of the ‘true quality’ of each student’s writing over the course of a semester – nevertheless that grade would seldom turn up on the transcript. That is, when a transcript says that a student got a B– for a writing course, we have no way of knowing whether it means: “B– is my judgment of the quality of the student’s writing for the semester”; or “He turned in consistently A quality writing but it was almost always late, and he cut lots and lots of classes and did not participate consistently with other class activities”; or “His writing was fairly weak, but he improved his writing enormously over the course of the semester – and he was super diligent about everything, even to the point of doing lots of extra work.”

My argument, then is that most course grades are not – and cannot be – fair representations of the quality of students’ writing. Given that fact, there is no need for a single standard or rubric or set of criteria for quality, even for course grades; multiple criteria are no impediment.

When I talk this way, I seem to be taking an anarchic position that says, “Course grades based on a single standard are unfair and so we might as well give up grading altogether.” This, in fact, is my belief – and it is a belief I could live by during my 9 years of teaching at The Evergreen State College where we used no numerical grades. But most of my teaching has been in institutions with conventional grading. Still, in recent decades, I have discovered that I can produce adequate one-dimensional course grades that are *not* based on judgments of the quality of student writing. I have worked out a form of *contract grading*. It fits perfectly with the theory of valuing that I have been laying out in this essay. It would also fit well with Inoue’s experiment of giving students the power to grade – though up till now I have never thought to (dared to?) hand over power to students as he does. Moreover, I find that this contract improves the conditions for teaching and learning and pretty consistently satisfies students as to fairness.

The contract is perhaps an unlovely hybrid affair: it decouples judgments of quality from grades – but only up through the grade of B. That is, any student who fulfills all the tasks required by the terms of the contract is guaranteed a course grade of B, irrespective of any of my judgments of writing quality. As for grades higher than B, I give them only when I judge that the students’ writing (especially their final portfolio) is exceptionally excellent. In this system, students get judgmental feedback and evaluation on their writing consistently throughout the semester from me and fellow students, but grades are not involved in these conversations. That is, *decoupling* evaluation does not mean *jettisoning* evaluation.

Grades up to a B, then, are linked to *activities*, not to judgments of quality. The list of tasks required for a B makes students spend 14 weeks engaging in all the activities and processes that I have found to be the most reliable causes of learning. In effect, the contract requires students to do everything I previously wished I could get them to do by means of a regular grading system: lots of writing (low stakes and high stakes), lots of revising, a good deal of copy editing, lots of giving and getting peer feedback, meeting almost all deadlines, showing up for almost all classes,

and doing all the class activities. (See [Appendix A](#) for the full list.) Students get lower grades when they fail to do some of these things. (I have used this kind of contract grading in first year writing classes for at least a dozen years, and other versions for other courses too. When I was writing program administrator, I offered contract grading as an option to graduate instructors. Jane Danielewicz is another director of a large writing program who has found the contract effective. (See [Danielewicz and Elbow \[2006\]](#) for fuller explorations of this kind of contract grading.)

Why the hybridity? It is the blunt compromise I make because I am teaching in an institution committed to grades. I have found over the years that if I make students spend 14 weeks doing all the things that most reliably produce learning, their writing actually ends up good enough to deserve the B I promised them (especially given the fact that B is such an ambiguous grade). I would use a contract for an A if I could figure out 14 weeks worth of process-requirements that would invariably lead to writing that is good enough so that I feel comfortable promising an A. I have not been able to manage that.

Because of this problem, I am stuck with a process for giving A grades that is guilty of all the sins I argue against in this essay. It is galling. My only defense is crude but powerful: the hybrid system lets me give *far fewer* grades that are irresponsible and intellectually incoherent than I used to give. That is, I am stuck having to give relatively few grades based on judgments of the quality of students' writing. So even though my very high grades are problematic for being based on my single, unilateral standard of good writing, I do not need to award so many of these grades. I give them only to students who fulfill these two conditions: (a) they fulfilled all the terms of the contract; (b) I judge that a great deal of their writing – especially their portfolio – is exceptionally excellent. At the end of the semester, there are usually a number of very skilled writers who dropped the ball in one way or another and did not fulfill the contract. I do not even have to think about the quality of their writing any more: I have already given them plenty of evaluative response throughout the semester. (Interestingly, I find this system helps a bit with grade inflation. It makes higher grades extremely 'special'; I can keep reminding students to look in the catalogue where it says that "B is an honors grade.")

In effect, my very high grades are no different than those produced by the conventional system (and thus no less invalid and unreliable). But I could make them *more* valid and reliable if I grafted my contract grading onto my revised version of Inoue's course (where students would not have to agree on a single standard). I have never done this, but if I did, I might come up with a formula saying something like this: students get grades of B+, A–, and A, respectively, if 60%, 70%, and 80% of the class voted their writing to be exceptionally excellent. Note the benefits. (1) These high grades would not be based on the verdict of a single person but rather of multiple minds. (2) They would be based on standards that were worked out gradually over time through the impressive social process that Inoue set up. Students cannot just say that X is better than Y, but must articulate their criteria and publicly compare and negotiate them. The crucial theoretical and practical point is that we can ask students to compare and negotiate and even debate – *without* forcing them to reach agreement. (3) These high grades would not be based on a single standard – thus avoiding the invalid pretense that there is one "true value" for "good writing." Some of the standards or criteria would represent group agreements; others would be idiosyncratic and wielded by just one or two students. Remember too, that *all* students throughout the term would be getting feedback from other students based on these multiple criteria.

Nevertheless, these improved higher grades would still be subject to one crucial impediment that I mentioned earlier. They would still be one-dimensional attempts to represent the quality of

multidimensional phenomena – and thus less than valid and reliable. For different graders (and readers of those grades) would give differential weightings to different dimensions of the writing, depending on their sense of the construct ‘good writing’.

I am willing to argue (but only gingerly) that this problem is *slightly* diminished in the case of higher grades. That is, a one-dimensional judgment of ‘excellent’ is probably a *bit* more defensible than a one-dimensional judgment of B and C and D. Middle range papers almost always represent a mixture of strengths and weaknesses, and so disagreements about middle range grades stem from conflicting priorities (for example between form and content). But when writing strikes readers as truly outstanding, there is at least greater likelihood that more dimensions are very good. Over the years, I *think* I see a bit more reader agreement about very excellent writing, and more reader disagreement about middle range grades. (Liz Hamp Lyons notes a number of studies that have noted better reliability with high (and low) scores – often as a byproduct of other conclusions. She cites Myford, Marr, and Linacre [1996] as one example [Personal communication].)

This admittedly weak defense of one-dimensional grades at the high end becomes a *bit* stronger when we use Inoue’s system. That is, if a large majority of the class, using multiple criteria, judges a portfolio to be truly outstanding, we are probably a little more justified in calling the writing indeed outstanding.⁷

7. Back to the main theme

In closing, I want to return to the main theme of my essay – value in writing – and link it to the question of why I celebrate such an ungainly hybrid version of contract grading. It is because educational institutions have succumbed to the widespread and pernicious assumption that Inoue fell into and that I am fighting here: the assumption that one-dimensional quantitative grades can

⁷ I am attracted to the idea of trying Inoue’s experiment in giving students power to determine criteria and grades. But I wouldn’t hand over decisions about which activities are required for the B. I see this as a *pedagogical* or *curricular* decision about which teaching-and-learning activities are most reliable for producing learning — not so much an assessment decision. In short, I would keep the contract unilateral rather than negotiated. Nevertheless, I retain the word “contract” because as teacher I’m making a written contractual promise: I guarantee them a B if they do all the requisite activities. To see examples of three teachers who use a form of contract grading where they *do* negotiate with students about which activities to require, see Shor (1996), Thelin (2005), and Moreno-Lopez (2005). The negotiations they describe strike me as vexed and they take too much class time away from the teaching of writing. Contract grading has done more to increase my satisfaction in teaching in an institutional setting than any other technique or approach I have used. Here are some of the ways it has improved the conditions for learning and teaching:

- The contract vastly reduces the amount of student resistance and anxiety about grades. When students get low grades, they can’t say, “She just doesn’t like me” or “She has a weird idea of good writing” — they know that low grades come only from things like late or missing papers.
- Teachers who fear that they will have no grades lower than a B can put their minds at ease. Plenty of students miss too many classes or fail to make substantive revisions — especially first year students.
- Teacher evaluations become significantly more effective for learning when they are decoupled from grades. Under normal grading conditions, students pretty much have to revise in accordance with the teacher’s evaluations and suggestions. That is, if the teacher says the organization is weak and makes a suggestion, students will assume they need to go along. They may not like the evaluation or feedback, but too seldom do they really think hard and ponder: “Is the organization really weak? Is that suggestion right? Could it be that there’s no problem? Or could it be that there’s some other problem that caused the teacher to talk about organization?” But that’s just what the contract invites them to do. Yes, my authority (in the contract) forces them to make substantive revisions, but they know that my evaluation of quality is entirely irrelevant to grades up to B. They can do something entirely different from what I suggest. They are, at long last, put in a position of having to think about the validity of response and evaluation.

fairly be based on a single standard and can be a valid representation of the quality of writing — which in turn rests on the assumption that there is a single standard of good writing.

It might seem as though institutions would be plunged into anarchy if they gave up that assumption — i.e., they would have to give up grading altogether or give only discursive evaluations (as at Evergreen State College, Hampshire College, and a number of others). But in fact, educational institutions could make a huge improvement on conventional grading without having to go that far — though the following suggestion represents another foray into compromise or hybridity.

Institutions could issue transcripts in which one-dimensional grades are replaced by multi-dimensional grids. For example, students might get grades of 1–4 on a number of criteria. The criteria would vary depending on the discipline. (It would be possible, though sad, to get a simple crude list for all courses such as “understanding”, “thinking”, “diligence.”) Teachers of writing could use a simple traditional set of criteria (like “Ideas”, “Organization”, “Clarity”, “Mechanics”, “Diligence”) or a much more sophisticated set. The numerical grades on any such rubric would still, of course, imply a single standard of good writing, but the damage would be considerably reduced on both theoretical and practical grounds. For there would be no single one-dimensional grade that vainly tries to represent the quality of a multidimensional performance. That bottom line judgment would have to be made by the *readers* of the grid. Different readers would give different weightings to different criteria in the transcript grid. For example, some readers would rate ‘mechanics’ very high; others not. In this way, different readers would be wielding different ‘constructs’ for ‘good writing’.

In effect, this multidimensional final grade would contain the *potentiality* of multiple standards for ‘good writing’. This would not do full justice the multiplicity of meanings in the world for good writing, but it would be a big improvement over present practice. And let us remember that any decent grid would have to have at least one item like “Diligence” or “Meets class responsibilities.” Many readers of transcripts see this as more central to the construct of ‘successful student’ (or good employee) than ‘organization of writing’. If this proposal for a change in college transcripts sounds too extreme, note that many or most teachers in the early grades give multidimensional final grades of exactly this sort. In doing so, they are using a wiser system — wiser because it enacts the recognition that we should avoid trying to legislate a single one-dimensional construct for goodness’ in second grade English or third grade social studies.

Appendix A

Main points in the grading contract as I have recently used it:

A.1. You are guaranteed a final grade of B if you meet the following conditions:

- (1) *Attendance.* Do not miss more than 1 week’s worth of classes. (If you miss class, you still need to do the assignment.)
- (2) *Lateness.* Do not be habitually late. (If you are late or miss a class, you still are responsible to find out what assignments were made.)
- (3) *Late assignments.* Do not have more than one late major assignment and one late smaller assignment.
- (4) *Journals.* Keep up your journal assignments.
- (5) *Sharing and responding.* Work cooperatively in groups. Be willing to share some of your writing, to listen supportively to the writing of others and, when called for, give full and thoughtful responses.

(6) Major assignments need to meet the following conditions:

- Always include process letter, all previous notes and drafts, and all feedback you got.
- *Revisions*. When the assignment is to revise, do not just correct or touch up. Your revision needs to reshape or extend or complicate or substantially clarify your ideas – or relate your ideas to new things. Revisions do not have to be better, but they must be different.
- *Copy editing*. When the assignment is for the final publication draft, your paper must be well copy edited – that is, free from virtually all mistakes in spelling and grammar. It is fine to get help in copy editing. (Copy editing does not matter for early and mid-process drafts.)
- *Perplexity*. For every paper, you need to find some genuine question or perplexity. That is, do not just tell four obvious reasons why dishonesty is bad or why democracy is good. Root your paper in a felt question about honesty or democracy – a problem or an itch that itches you. (By the way, this is a crucial skill to learn for success in college: how to find a question that interests you – even in a boring assignment.)
- *Thinking*. Having found a perplexity, then use your paper to do some figuring-out. Make some intellectual gears turn. Thus your paper needs to move or go somewhere – needs to have a line of thinking.
- Do not let these last two conditions bother you. I do not ask that your essays always be tidy, well organized, and perfectly unified. I care more about working through the question than about finding a neat answer. It is okay if your essays have some loose ends, some signs of struggle – especially in early drafts. But lack of unity or neatness needs to reflect effort, not lack of effort.

A.2. *Getting an A/B or A*

As you see, the grade of B depends on behaviors. Grades of A or A/B, however, depend on quality. Thus you earn a B if you put in good time and effort; I will push you all to get a B. But to get an A or A/B, you have to make your time and effort pay off into writing of genuine excellence (and also meet the conditions for a B). Notice that for grades up to B, you do not have to worry about my judgment or my standards of excellence; for higher grades you do. But we will have class discussions about excellence in writing and usually we can reach fairly good agreement. Your mid-semester and final portfolios will play a big role in decisions about excellence.

A.3. *Knowing where you stand*

This system is better than regular grading for giving you a clear idea of what your final grade looks like at any moment. For whenever I give you feedback on any major assignment, I will tell you clearly if you have somehow failed to satisfy the contract for a B. I will also tell you if I judge your draft to be genuinely excellent and thus to exceed the contract for a B. As for absences and lateness, you will have to keep track of them, but you can check with me any time.

A.4. *Grades lower than B*

I hope no one will aim for these grades. The quickest way to slide to a C, D, or F is to miss classes and show up without assignments. *This much is non-negotiable: you are not eligible for a passing grade of D unless you have attended at least 11 of the 14 weeks worth of classes, and completed 90% of the assignments. And you cannot just turn in all the late work at the end.* If

you are missing classes and behind in work, please stay in touch with me about your chances of passing the course.

References

- Bizzell, P. (2002). The intellectual work of 'mixed' forms of academic discourse. In: C. Schroeder, H. Fox, & P. Bizzell (Eds.), *Alt dis: Alternative discourses and the academy* (pp. 1–10). Portsmouth, NH: Boynton/Cook Heinemann Press.
- Broad, B. (2003). *What we really value: Beyond rubrics in teaching and assessing writing*. Logan: Utah State University Press.
- Danielewicz, J., and Elbow, P., (2006). *A unilateral grading contract to improve the conditions for teaching and learning*. Manuscript submitted for publication.
- Elbow, P. (1973). Appendix essay. The doubting game and the believing game: An analysis of the intellectual process. In: *Writing without teachers* (pp. 147–191). New York: Oxford University Press.
- Elbow, P. (1986). Methodological doubting and believing: Contraries in inquiry. In: *Embracing contraries: Explorations in learning and teaching* (pp. 254–300). New York: Oxford University Press.
- Elbow, P. (1991). Foreword. In: P. Belanoff & M. Dickson (Eds.), *Portfolios: Process and product* (pp. ix–ix10). Portsmouth NH: Boynton/Cook Heinemann Press.
- Elbow, P. (2005). Bringing the rhetoric of assent and the believing game together — and into the classroom. *College English*, 67, 388–399.
- Gardner, H. (1956). *Limits of literary criticism: Reflections on the interpretation of poetry and scripture*. London: Oxford University Press.
- Herrnstein Smith, B. (1988). *Contingencies of value: Alternative perspectives for critical theory*. Cambridge: Harvard University Press.
- Hirsch, E. D., Jr. (1977). *The philosophy of composition*. Chicago: University of Chicago Press.
- Inoue, A. B. (2005). Community-based assessment pedagogy. *Assessing Writing: An International Journal*, 3, 208–238.
- Moreno-Lopez, I., (2005). Sharing power with students: The critical language classroom. *Radical Pedagogy*, 7 (2). Retrieved from http://radicalpedagogy.icaap.org/content/issue7_2/moreno.html.
- Myford, C., Marr, D. B., & Linacre, J. M. (1996). *Reader calibration and its potential role in equating for the Test of Written English. TOEFL Research Report 52*. Princeton, NJ: Educational Testing Service.
- Newkirk, T. (1997). *The performance of self in student writing*. Portsmouth: Boynton/Cook Heinemann.
- Slomp, D. H., & Fuite, J. (2005). Following Phaedrus: Alternate choices in surmounting the reliability/validity dilemma. *Assessing Writing: An International Journal*, 3, 190–207.
- Shor, I. (1996). *When students have power: Negotiating authority in a critical pedagogy*. Chicago: University of Chicago Press.
- Smith, W. (1992). The importance of teacher knowledge in college composition placement testing. In: J. R. Hayes, R. E. Young, M. L. Matchett, M. McCaffrey, C. Cochran, & T. Hajduk (Eds.), *Reading empirical research studies: A rhetoric of research* (pp. 287–316). Hillsdale NJ: Erlbaum.
- Thelin, W. H. (2005). Understanding problems in critical classrooms. *College Composition and Communication*, 57, 114–141.