Introduction: Degrees of Freedom in Assessment, Evaluation, and Grading

Stephen Tchudi
Chair, NCTE Committee on Alternatives to Grading Student Writing

One of the highlights of my week is a trip to the neighborhood grocery/deli with my fourteen-year-old son Chris, to indulge in a couple of doughnuts and a shared reading of the *Weekly World News*. For those unfamiliar with this tabloid, you'll find it at the checkout counter next to the *National Enquirer* and the *Star*, papers that Chris and I regard as mere scandal sheets. *Weekly World News*, to our way of thinking, is more serious journalism—though not in the league of the *Washington Post*. The *News* entertains us with stories of possibly true and certainly odd happenings from around the globe.

For example, we read of a math teacher from Herford, Germany, who was "fired for making kids eat night crawlers!" (30 Jan. 1996: 4). It seems that William Enbeck used "sadistic" assignments to let students raise their grades. He gave a girl thirty points "for pushing a peanut around the classroom with her nose for an hour"; he awarded a passing grade to the lad who "submerged his head in a bucket of cold water for 45 seconds." Another kid went up the grading scale while eating "half a pound of long, fat earthworms." At the hearing called by irate parents and administrators, Enbeck defended himself by claiming that these were character-building exercises that demonstrated students' "guts and willingness to take risks—qualities that will take them much farther in life than basic math."

As Chair of NCTE's Committee on Alternatives to Grading Student Writing, I have to confess that I found Enbeck's alternatives quite interesting. He articulated clear goals; he set specific assignments with obvious criteria for evaluation; and if the students did the work, they got the reward. His grading scheme seemed a lot simpler and in some ways no more arbitrary and subjective than those devised by many of us more conventional elementary, secondary, and college writing

teachers. Now, I don't think I'll be having my students eat worms, though some might say that my assignments are just as unpleasant. But hyperbole aside, the Enbeck story dramatizes the dilemma of grading that prompts this book: the arbitrariness of grades, the use of grades to coerce students into performance, and the irrelevance of grades to the sort of authentic assessment one experiences in life.

My personal quest for alternatives to grading student writing began more than thirty years ago, during my first year of college teaching. At Northwestern University, I was teaching an undergraduate course called "Practical Rhetoric." In the spirit of the then recent Dartmouth Conference, I had turned this traditionally analytic course into a writing workshop with an emphasis on the process of composing, from prewriting through editing. I graded the first round of papers by conventional standards: C = OK, B = Better, A = Excellent. I had no complaints from the students, who mostly received B's and A's, thanks to the possibility of revision through writing workshops. But for the second round of papers, I hit a snag. Student Julie, who had received an A on her first paper, complained about the B I gave her this time. I explained and justified the grade to her: I thought that the second paper lacked the verve and voice of the first and that it showed some signs of hasty revision. In words that are echoed in an essay by Jean Ketter and Judith Hunter elsewhere in this collection, Julie said, in effect, *"This is the best I can do; it's my very best!"* I suggested that she might do further revision, but I might as well have asked her to eat night crawlers. She did C-level work for the rest of the course and seldom talked to me. This experience was something of an epiphany for me. To this day, I blame the grading system for poisoning my teacher/student relationship with Julie, and since that course, I have never again put a letter grade on a piece of student writing.

I have sought alternatives. For a while, I used a "recommended" grading system, where students would do self-assessment and justify or argue for a grade. The scheme left me uneasy, however, especially when a colleague accused me of making the students do what I, myself, was unwilling to do: reduce a range of achievements to a letter grade. I've tried a variety of contract- and performance-based systems; I've worked in pass/fail and nongraded systems; and I've used holistic as well as analytic scoring of portfolios. Each of these schemes offered me certain degrees of freedom to assess and respond to student writing, and each had certain drawbacks, restrictions, and problems. Always problems.

Each semester, as I'm preparing my syllabi, I agonize over the grading criteria. What will it be this semester? Contracts? Achievement grading? A point system? Portfolios? Negotiated grades? Self-evaluation? In what combination? I find myself in the absurd position of rearranging course content in order to accommodate the demands of the grading system. I've been known to develop several alternative syllabi, carefully saved in the computer, and to make my final selection of a grading system only at the last minute, when forced to photocopy the syllabus for those students who will show up an hour later and want to know, "How do I get an A out of this course?" or "How can I get out of this course with a C?"

I'm almost phobic about this aspect of my teaching, and that's why I accepted the invitation to chair the NCTE Committee on Alternatives to Grading Student Writing. We were charged by the Executive Committee of NCTE

> to investigate all alternatives to giving students grades in writing so that progress can be evaluated in ways sensitive to the needs of students as well as universities, colleges, and school districts; to organize the results of that investigation through manuscripts that help teachers and others in elementary, middle, and secondary schools and in colleges and universities to understand the theory and practice of alternatives to grading; to set a timely schedule for the gathering of information and the submitting of a prospectus and manuscript to the NCTE Editorial Board.

I saw chairing this committee as an opportunity to learn new approaches and alternatives, and maybe even to find pedagogical salvation: the Perfect Grading System that would be true to the research on grading, consistent with current writing pedagogy, fair to students, and productive in moving students toward being highly motivated, highly skilled writers.

Well, the committee *hasn't* found that *perfect* system, although the readers of this book will learn about some very powerful alternatives to grading student writing. I have personally learned a great deal from the members of the committee (listed in the front matter of this book) and from the contributors to this volume. I have experimented with the grading schemes described here and have found that they produced a good deal of peace of mind, though not complete salvation. Actually, none of the writers has found or claims to have found that perfect system, but everyone represented in this book is working at it constantly, as, I suspect, are the readers who pick up this volume.

We're a community, we searchers for alternatives to grading student writing.

The committee is convinced by the research presented in Part I of this book, which shows quite clearly that grading writing doesn't contribute much to learning to write and is in conflict with the new paradigms for writing instruction. As a committee, we would unanimously love to see grades disappear from education altogether so that teachers and students can focus on authentic assessment, but we realize that in the current educational climate, that's not likely to happen. Although a few schools and colleges are experimenting with nongraded systems, and although a growing number of school systems do not grade the youngest students, the vast majority of English/language arts teachers will, throughout their teaching careers, be faced with the periodic need to sum up students' work by some set of criteria and to translate that performance into a grade that goes on a report card or transcript. From the outset, we wanted this book to be theoretically sound but classroom practical, supplying genuine alternatives for teachers who work under the shadow of grading systems.

We think it is useful to conceive of the problem by adapting a concept from math and science: "degrees of freedom." Both numbers and molecules have constraints on their freedom to move, to vary. Changing parameters or restrictions often opens up new areas of freedom, but just as often results in the loss of other directions of movement. So it is, we think, with grading. The aim of this book is to help teachers increase their freedom to explore alternatives in assessment.

Figure 1 shows our interpretation of this concept and will help the reader understand the structure and philosophy of this book. The figure represents the tension between what research and teacher instinct tell us—to broaden the range of possibilities for assessment and reaction to student work—versus the pressures to place a single grade/symbol on the final product.

The committee thinks it important to distinguish among four interrelated terms and concepts shown in Figure 1: *response, assessment, evaluation,* and *grading*. From top to bottom, these represent *decreasing* degrees of freedom in reacting to and evaluating student work. Response to writing is, we believe, at the heart of the process. As Lynn Holaday says in her essay in this volume, "*Writing students need coaches, not judges.*" *Response* is so important, in fact, that the committee wanted an entire section of the book (Part II) devoted to response strategies, without any reference to grades. Response to writing has the greatest range of freedom because it is naturalistic, growing direct-

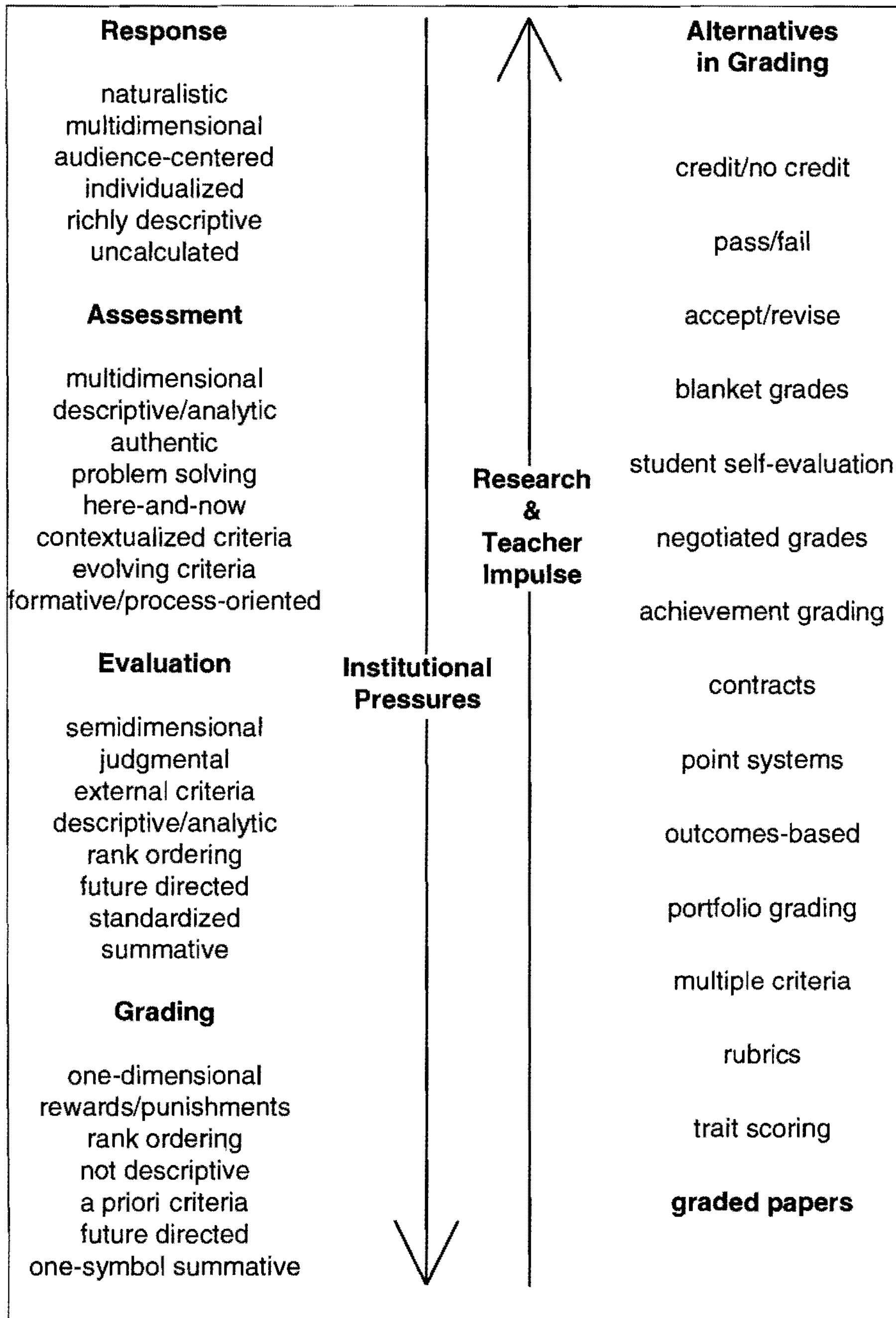

Figure 1. Degrees of freedom in assessment, evaluation, and grading

ly from readers' reaction to a text. Somebody loves it; somebody else hates it; a third person wants to punch the writer in the nose.

Although most teachers agree that "real" response from "real" audiences is desirable, there are some reasons to make certain that response is productive and civil (unlike a punch in the nose), which is why we teach writing in classrooms rather than simply letting novices pick it up through the school of hard knocks. The essays in Part II, then, suggest ways in which, without unduly limiting freedom, the teacher can make response to writing an increasingly productive part of the writing process.

Moving down the left-hand column of Figure 1, we are also persuaded that there is a need for what we'll call "instructionally calculated" reaction to student writing—peers, teachers, parents, or others saying things to writers that will help them write better.

Assessment in writing is a bit more limited in its degrees of freedom than response. Assessment asks broadly: "How did this project turn out?" "How is it turning out?" "Are you getting what you want from your audience?" "Can I suggest a few possibilities that might make this thing work better?" "What do you think your work needs?" "What are your ideas about how to make it more successful?" Assessment certainly incorporates reader response (to drafts as well as finished products), and it is often focused on practical, functional concerns: "What do I have to do to make this paper *work?*" Frequently, criteria for assessment evolve *as* one writes: "How do I know what I think until I see what I say?" The purpose of assessment is very much in the here and now, rather than in the future; that is, the concern is not so much "How do I get to be a better writer in the future?" as "How can I be successful with this paper right now?" Assessment uses a lot of description (rather than judgments) of readers and writers and is "formative" or in-process rather than "summative" or final.

We think it is especially important to distinguish between assessment and the next level, *evaluation* of writing. Too often, these terms are used indiscriminately and are sometimes even conflated with "grading." I regularly hear teachers say, for example, that "I have to evaluate some papers," when they will, in fact, be grading them. Or, as the committee heard on several occasions, people sometimes falsely assume that if you're opposed to grading writing, you're refusing to evaluate or assess, just accepting any old thing a student writes. For clarity of discussion, then, it's important to say "assessment is not evaluation is not grading," although the three are clearly linked.

To distinguish from the broader term, assessment, we argue that *evaluation* implies fixed or a priori criteria rather than evolutionary or constructed values. To "evaluate" means to compare work with some sort of marker, benchmark, or standard. Unfortunately, many students see evaluation as essentially punitive: "No, James, this paper is not up to par." However, the best evaluation can be constructive: "I think this paper moves beyond what you were doing in your last one" or "Yes, I think this paper is just about ready for publication." Where assessment criteria tend to be developed within the rhetorical context of a project, evaluation looks to established criteria: "This paper does/doesn't have a clear sense of purpose" or "The language is/is not gobbledygook" or "Your spelling is/is not up to standard." Evaluation also tends to be comparative—using phrases like "better than" or "worse than"—and it lends itself to rank ordering: "top to bottom," "upper third," or "90th percentile."

Both evaluation and assessment are a natural part of the writing process; writers assess and evaluate constantly, from idea to printed page. The committee (supported by the new paradigm in writing) believes that assessment, along with cultivated response, is the most useful kind of information that writers can receive. Evaluation, although more restricted in the range of commentary it can offer, seems a natural enough element in writing, for we all want to know "How'm I doin'?" and "How can I do better?"

But then there is *grading*, the fourth term in our model. As Leisel O'Hagan shows us in the next essay, very few people have anything nice to say about grading. Grades are extremely limited in their degrees of freedom, for they take a vast array of data and condense it into a single symbol that, in itself, doesn't communicative very much. Grades are one-dimensional, and they tend to be based on a priori, even Platonic, notions of "good" and "bad." As a result, grades stereotype, pigeonhole, and rank order students and their writing. Any grade less than an A destroys student morale to some extent, and even the prized A falsely implies that the student has reached a kind of perfection. Above all, grades fail to provide descriptive information of any significance—what, after all, does a C or an A tell you about how people responded to your work or even what you might do with the next piece you write? Furthermore, grades reduce students' degrees of freedom to internalize advice about writing: Why should you pay attention to what the teacher or peers said when all that matters is that grade?

The focus of Part III is on practical alternatives to grading student writing, systems and techniques that maintain many degrees of freedom in reacting to students' work while still satisfying the institutional need to derive a grade. As summarized in the right-hand column of Figure 1, the papers in the collection explore systems ranging from pass/fail (with very high degrees of freedom) to rubrics (which increase degrees of freedom over traditional grades though careful delineation of criteria). In between, the reader will find essays that deal with such approaches as collaborative evaluation, contract grading, achievement grading, outcomes-based assessment, portfolio grading, and total quality evaluation.

In addition, sprinkled among these "how to" essays, the reader will also find some statements we call "Interludes," gleaned from the committee's discussion folder on NCTENet. We think these brief statements offer some provocative thoughts about grading issues and offer further techniques and strategies being tried effectively by teachers all across North America.

Part IV closes the book with a set of outlines for faculty workshops in alternatives to grading student writing. Although virtually all of the grading alternatives described in Part III can be done by the individual teacher in his or her own classroom, there is both safety and the possibility of increased curriculum coherence in numbers. If grading alternatives are to achieve their true potential for helping students and teachers think differently about writing, response, assessment, and evaluation, faculties need to work together systematically to develop, implement, and evaluate new approaches. The workshops are designed with that in mind and are intended for use by small clusters of like-minded teachers or by entire school buildings or districts. The workshops were developed by members of the NCTE Committee on Alternatives to Grading Student Writing and were field-tested at the NCTE Annual Convention in San Diego.

Finally, I want to close this introduction by appealing briefly to potential readers of this book who may *not* be sold on the whole concept of alternatives to grading student writing or who are in systems that specifically require letter grades to be placed on individual papers. In our work as a committee, we have heard from articulate teachers who argued that the problem is not with grading, but with *how* teachers grade. We have also heard from teachers who would like not to use grades but explain that "the principal has mandated it." The committee members want to share our belief that even in such cases, there are bad, good, and better ways of going about grading writing—

fair and unfair; equitable and inequitable; destructive, constructive, and highly constructive. There's certainly no support in the professional literature for arbitrary grading of the "guess what the teacher wants this time" variety. There is no support for punitive grades or grades used as a way of coercing students into studying material that is too difficult or irrelevant. Moreover, there *is* support for the processes of evaluation and assessment described in this book: involving students collaboratively in the response and assessment; using carefully articulated criteria; rewarding growth and effort as well as perceived absolute quality; including large amounts of response to student writing; seeking real-world audiences; and focusing writing as a whole on issues, topics, and concerns that are important to the writer. Thus, a reading of this book will, we hope, help the teacher who uses conventional grading to make that approach potentially and pedagogically more useful.

Still, the NCTE Committee on Alternatives to Grading Student Writing finds that both teacher experience and educational research argue powerfully for the abolition of letter grades on individual student papers. *We prefer and promote alternatives to grading student writing.*