

RETHINKING RUBRICS IN WRITING ASSESSMENT

MAJA WILSON

HEINEMANN
PORTSMOUTH, NH

Contents

Acknowledgments	ix
Foreword by Alfie Kohn	xi
Introduction: When Best Practice and Our Deepest Convictions Are at Odds	xix
1 My Troubles with Rubrics	1
2 There Is a Cow in Our Classroom: How Rubrics Became Writing Assessment's Sacred Cow	10
3 The Broken Promises of Rubrics	27
4 The Golden Rule of Assessment: Why We Don't Practice for Assessment What We Preach for Pedagogy	43
5 Agreeing to Disagree: The Heart of a New Writing Assessment Paradigm	53
6 Making Our Subjectivity Transparent and Useful: What Response Unmediated by Rubrics Looks Like in Our Classrooms	66
7 But How Shall We Grade? Investing in Process for the Sake of Product	79
8 How Do We Make Time to Make It Meaningful? And Other Questions About Assessment Without Rubrics	88
References	99
Index	103

Foreword

by Alfie Kohn

Once upon a time I vaguely thought of assessment in dichotomous terms: The old approach, which consisted mostly of letter grades, was crude and uninformative, while the new approach, which included things like portfolios and rubrics, was detailed and authentic. Only much later did I look more carefully at the individual floats rolling by in the alternative assessment parade—and stop cheering.

For starters, I realized that it's hardly sufficient to recommend a given approach on the basis of its being better than old-fashioned report cards. By that criterion, just about anything would look good. I eventually came to understand that not all alternative assessments are authentic. My growing doubts about rubrics in particular were prompted by the assumptions on which this technique rested and also the criteria by which they (and assessment itself) were typically judged. These doubts were stoked not only by murmurs of dissent I heard from thoughtful educators¹ but by the case made *for* this technique by its enthusiastic proponents. For example, I read in one article that "rubrics make assessing student work quick and efficient, and they help teachers to justify to parents and others the grades that they assign to students."²To which the only appropriate response is: Uh-oh.

I had been looking for an alternative to grades, and the reason I was keen to find one is that research shows three reliable effects when students are graded: They tend to think less deeply, avoid taking risks, and lose interest in the learning itself.³ The ultimate goal of authentic assessment must be the elimination of grades. Obviously, a strategy that merely offered a new way to arrive at those final marks wouldn't address the fundamental problem of students who had been led to focus on getting As (or their equivalent) rather than on making sense of ideas. Moreover, something that was commended to teachers as a handy strategy of self-justification during parent conferences ("Look at all these 3s, Mrs. Grommet! How could I have given Zach anything but a B?") didn't seem particularly promising for inviting teachers to rethink their practices and premises.

As for the selling point of "quick and efficient," I've graded enough student papers to understand the appeal of this promise. Still, the best

teachers would react with skepticism, if not disdain. They'd immediately ask what we had to sacrifice in order to spit out a series of tidy judgments about the quality of student learning. To ponder that question is to understand how something that presents itself as an innocuous scoring guide can be so profoundly wrongheaded.

The first problem is that the how's of assessment (when they preoccupy us) chase the why's back into the shadows. So let's shine a light over there and ask: What's our reason for trying to evaluate the quality of students' efforts? This is a question we rarely ask, but it matters whether the objective is to (1) rank kids against one another, (2) provide an extrinsic inducement for them to try harder, or (3) offer feedback that will help them become more adept at, and excited about, what they're doing. I worry that giving teachers more efficient rating techniques—and imparting a scientific luster to those ratings—may make it even easier to avoid asking this question. In any case, it's certainly not going to shift our rationale away from (1) or (2) and toward (3).

Second, consistent and uniform standards are admirable, and maybe even workable, when we're talking about, say, the manufacture of DVD players. The process of trying to gauge children's understanding of ideas is a very different matter, however, and ought to be treated as such. It necessarily entails the exercise of human judgment, which is a messy, imprecise, subjective affair. Rubrics are, above all, a tool to promote standardization, to turn teachers into grading machines or at least allow them to pretend that what they're doing is efficient, exact, and objective. Frankly, I'm amazed by the number of educators whose opposition to standardized tests and standardized curricula mysteriously fails to extend to standardized in-class assessments.

The appeal of rubrics is supposed to be their high interrater reliability, finally delivered to language arts—the “transformation of English classes into something as rigorous and legitimate as biology or chemistry classes,” as Maja Wilson puts it. A list of criteria for what should be awarded the highest possible score when evaluating an essay is supposed to reflect near-unanimity on the part of the people who designed the rubric and is supposed to assist all those who use it to figure out (that is, to discover rather than to decide) which essays meet those criteria.

Now some observers criticize rubrics because they can never deliver the promised precision; judgments ultimately turn on adjectives that are murky and end up being left to the teacher's discretion. But I worry more about the success of rubrics than their failure. Just as it's possible

to raise standardized test scores, providing that you're willing to gut the curriculum and turn the school into a test-preparation factory, it's possible to get a bunch of people to agree on what rating to give an assignment, providing that they're willing to accept and apply someone else's narrow criteria for what merits that rating. Once we check our judgment at the door, we can all learn to give a 4 to exactly the same things.

This attempt to deny the subjectivity of human judgment, this "fear of disagreement," as Wilson calls it, is objectionable in its own right. But it's also harmful in a very practical sense. In an important article published in 1999, Linda Mabry, now at Washington State University, pointed out that rubrics "are designed to function as scoring guidelines, but they also serve as arbiters of quality and agents of control" over what is taught and valued. Because "agreement among scorers is more easily achieved with regard to such matters as spelling and organization," these are the characteristics that will likely find favor in a rubricized classroom. Mabry cites research showing that "compliance with the rubric tended to yield higher scores but produced 'vacuous' writing."⁴

To this point, my objections assume only that teachers rely on rubrics to standardize the way they think about student assignments. Despite my misgivings, I can imagine a scenario where teachers benefit from consulting a rubric briefly in the early stages of designing a curriculum unit in order to think about various criteria by which to assess what students end up doing. As long as the rubric is only one of several sources, as long as it doesn't drive the instruction, it could conceivably play a constructive role.

But all bets are off if *students* are given the rubrics and asked to navigate by them. The proponent I quoted earlier, who boasted of efficient scoring and convenient self-justification, also wants us to employ these guides so that students will know ahead of time exactly how their projects will be evaluated. In support of this proposition, a girl who didn't like rubrics is quoted as complaining, "If you get something wrong, your teacher can prove you knew what you were supposed to do."⁵ Here we're invited to have a good laugh at this student's expense. The implication is that kids' dislike of these things proves their usefulness—a kind of "gotcha" justification.

Just as standardizing assessment for teachers may compromise the quality of teaching, so standardizing assessment for learners may

compromise the learning. Mindy Nathan, a Michigan teacher and former school board member told me that she began “resisting the rubric temptation” the day “one particularly uninterested student raised his hand and asked if I was going to give the class a rubric for this assignment.” She realized that her students, presumably grown accustomed to rubrics in other classrooms, now seemed “unable to function unless every required item is spelled out for them in a grid and assigned a point value. Worse than that,” she added, “they do not have confidence in their thinking or writing skills and seem unwilling to really take risks.”⁶

This is the sort of outcome that may not be noticed by an assessment specialist who is essentially a technician, in search of practices that yield data in ever-greater quantities. A B+ at the top of a paper tells a student very little about its quality, whereas a rubric provides more detailed information based on multiple criteria. Therefore, a rubric is a superior assessment.

The fatal flaw here is revealed by a line of research in educational psychology showing that students whose attention is relentlessly focused on how well they’re doing often become less engaged with *what* they’re doing. There’s a big difference between thinking about the content of a story one is reading (for example, trying to puzzle out why a character made a certain decision), and thinking about one’s own proficiency at reading. “Only extraordinary education is concerned with learning,” the writer Marilyn French once observed, whereas “most is concerned with achieving; and for young minds, these two are very nearly opposites.”⁷ In light of this distinction, it’s shortsighted to assume that an assessment technique is valuable in direct proportion to how much information it provides. At a minimum, this criterion misses too much.

But the news is even worse than that. Studies have shown that too much attention to the quality of one’s performance is associated with more superficial thinking, less interest in whatever one is doing, less perseverance in the face of failure, and a tendency to attribute results to innate ability and other factors thought to be beyond one’s control.⁸ To that extent, more detailed and frequent evaluations of a student’s accomplishments may be downright counterproductive. As one sixth grader put it, “The whole time I’m writing, I’m not thinking about what I’m saying or how I’m saying it. I’m worried about what grade the teacher will give me, even if she’s handed out a rubric. I’m more focused on being correct than on

being honest in my writing.”⁹ In many cases, the word *even* in that second sentence might be replaced with *especially*. But, in this respect at least, rubrics aren’t uniquely destructive. Any form of assessment that encourages students to keep asking, “How am I doing?” is likely to change how they look at themselves and at what they’re learning, usually for the worse.

The book you’re about to read is not only a superb analysis of rubrics but a lesson in how to apply careful thinking to classroom practice. There is an inviting informality to Wilson’s tone that manages to coexist with incisive analysis and careful organization. She’s persuasive by virtue of her arguments and experience, without a hint of pomposity—and she makes all this look effortless.

What really distinguishes Wilson’s analysis is her willingness to challenge rubrics not merely for their technical deficiencies but on the basis of the goals they serve. That’s a rarity in the world of assessment. She contends that improving the design of rubrics, or even inventing our own, will not suffice because there are problems inherent to the very idea of rubrics. She shows that their attempt to standardize assessment is rooted in an effort to rank students against one another—and she points out that neither we nor our assessment strategies can be simultaneously devoted to helping all students improve *and* to sorting them into winners and losers.

What seems to trouble Wilson most of all, though, is how rubrics are relentlessly reductive (a fact that can drive anyone to alarming alliteration). High scores on a list of criteria for good writing do not mean that what has been written is good, she explains, because quality is more than the sum of its rubricized parts. In fact, she suggests in passing that “we need to look to the piece of writing itself to suggest its own evaluative criteria”—a truly radical and provocative suggestion.

Wilson also makes the devastating observation that a relatively recent “shift in writing pedagogy has not translated into a shift in writing assessment.” Teachers are given much more sophisticated and progressive guidance nowadays about how to teach writing but are still told to pigeonhole the results, to quantify what can’t really be quantified. Thus, the dilemma, which she doesn’t shrink from describing, just as she doesn’t hesitate to identify by name who’s leading us astray: Either our instruction and our assessment remain “out of sync” or the instruction gets worse in order that students’ writing can be easily judged with the help of rubrics.

I love that Wilson is not only willing but able to think seriously about alternatives—that is, ways of evaluating writing that are grounded in her classroom experience and consistent with the values that inform her critique. The result is a book that I dearly hope will be read not only among people who teach writing (and think about teaching writing, and think about teaching writing to writing teachers), but among people whose specialty is assessment.

Rethinking Rubrics is, at its core, a rather incendiary piece of work. But then, as the late John Nicholls once remarked in response to my use of that word to describe another education book, “There’s a lot of trash to be burnt.”

Notes

-
- ¹ In addition to the sources I’m about to cite, I might mention critical comments about rubrics offered by Bruce Marlowe and Marilyn Page (in *Creating and Sustaining the Constructivist Classroom*, 2nd ed., [Thousand Oaks, CA: Corwin, 2005], p. 56); Susan Ohanian (in *What Happened to Recess and Why Are Our Children Struggling in Kindergarten?* [New York: McGraw-Hill, 2002], p. 40); Jackie Brooks (as quoted by Kathy Checkley, “Assessment That Serves Instruction,” *ASCD Education Update*, June 1997, p. 5); and Dennis Littky (as quoted by Eliot Levine, *One Kid at a Time* [New York: Teachers College Press, 2002], p. 116). Also, assessment specialist W. James Popham (“What’s Wrong—and What’s Right—with Rubrics,” *Educational Leadership*, October 1997, pp. 72–75) faults many rubrics for being overly detailed and for using criteria that are amorphous, if not tautological.
 - ² Heidi Goodrich Andrade, “Using Rubrics to Promote Thinking and Learning,” *Educational Leadership*, February 2000, p. 13.
 - ³ I review this research in *Punished by Rewards* (Boston: Houghton Mifflin, 1993) and *The Schools Our Children Deserve* (Boston: Houghton Mifflin, 1999), as well as in “From Degrading to De-Grading,” *High School Magazine*, March 1999 (available at www.alfiekohn.org/teaching/fdtd-g.htm).

- ⁴ Linda Mabry, "Writing to the Rubric," *Phi Delta Kappan*, May 1999, pp. 678, 676.
- ⁵ Quoted by Andrade, "Understanding Rubrics," in <http://learnweb.harvard.edu/alps/thinking/docs/rubricar.htm>. Another educator cites this same quotation and adds: "Reason enough to give rubrics a closer look!" It's also quoted on the RubiStar website, which is a sort of online rubric-o-matic.
- ⁶ Mindy Nathan, personal communication, October 26, 2004. As a student teacher, Nathan was disturbed to find that her performance, too, was evaluated by means of a rubric that offered a ready guide for evaluating instructional "competencies." In an essay written at the end of her student-teaching experience, she commented, "Of course, rubrics don't lie; they just don't tell the whole story. They crunch a semester of shared learning and love into a few squares on a sheet that can make or break a career." That's why she vowed, "I won't do this to my students. My goal as a teacher will be to preserve and present the human aspects of my students that defy rubric-ization."
- ⁷ Marilyn French, *Beyond Power: On Women, Men, and Morals* (New York: Summit, 1985), p. 387.
- ⁸ For more on the distinction between performance and learning—and the detrimental effects of an excessive focus on performance—see *The Schools Our Children Deserve*, chap. 2, which reviews research by Carol Dweck, Carole Ames, Carol Midgley, John Nicholls, and others.
- ⁹ Quoted in Natalia Perchemlides and Carolyn Coutant, "Growing Beyond Grades," *Educational Leadership*, October 2004, p. 54. Notice that this student is actually making two separate points. Even some critics of rubrics, who are familiar with the latter objection—that honesty may suffer when technical accuracy is overemphasized—seem to have missed the former one.

Introduction

When Best Practice and Our Deepest Convictions Are at Odds

More often than not . . . the field of composition has erred when it has too hastily trusted and laid claim to certainty. . . . Instead of providing solutions, the urge for certainty has often created new problems by encouraging simpleminded mechanical procedures for teaching or learning highly complex skills and processes. Guised in the cloak of reliability and efficiency, such procedures are instructionally very attractive, and teachers adopt them rapidly, often in spite of their deepest convictions about the complexities of the writing process.

—CHRIS ANSON (1989, 2)

Amy's Ford Pinto is packed with far too many clothes and coolers for a three-day weekend at her family's cottage, but that doesn't stop us from driving back to Jenny's house for the coffee maker, just in case. We'd hate to be stuck on a writing weekend without coffee. As Jenny climbs back into the car, balancing the coffee maker on the four laptops piled next to me, I mention how funny it would be for the math teachers in our district to get together over the summer for a math weekend. As we laugh at the image, Amy says in all sincerity, "We have the coolest subject."

Later that day, as Amy writes about nakedness and Laura curses Emily Dickinson and Jenny blurs the line between Harlequin Romance and personal narrative, I think about my deepest convictions about writing: about the magic of developing and finding a voice, about the constant struggle to meaningfully represent experience through words. I think about Frederick Douglass, whose literacy literally and metaphorically saved his life. I think about Maya Angelou, whose conviction about the power of words drove her to silence before it emerged in all the beauty of her writing. I think

about Emily Dickinson, who wouldn't go outside, but whose words have traveled through space and time and a certain slant of light. I think about my students, who wonder if words can carry the weight of their convictions and fears. And I think that Amy is right; English is the coolest subject.

Writing teachers are a passionate group. Our earliest and deepest experiences with language led us to this profession. We were seduced by the rhythm of language, or by the connection stories brought us with our parents, or by the way words allowed us to form and express our humanity. We were comforted by the way that writing anchored our thoughts on paper, allowing us to build solid ideas from fluid thoughts. We were amazed by the way that scribbles on paper could create understanding. We are convinced that there is something fundamentally sacred about teaching writing—about helping another person to express and shape their humanity through language.

Still, I think of the huge no man's land between my deeply held convictions about the power of writing and some of my classroom practices. I remember with a sinking feeling the quiz I gave in which I asked students to list the steps of the writing process and then define the word *recursive*. I remember the time Tim didn't know what to write next and I gave him an entire sentence instead of asking him questions to prompt his own words. I tell this to Amy in order to dispense with my guilt, and she admits to similar transgressions: the paper hastily graded, the writing starter quickly lifted from the book of canned journal prompts, the empty praise offered in a moment of weakness. We know that good teachers must constantly revisit their knowledge about best practice and question their performance in light of this knowledge, and we are comforted that at least we know when we fall short. It will take a lifetime to become good teachers, and we can accept our shortcomings as long as we are headed in the right direction.

What happens, however, when the field of writing methods leads us astray? For the most part, the things we learn from methods classes, *English Journal* articles, or professional conferences work well; no one would argue that using writing as punishment really works or that giving feedback only on spelling errors has been mistakenly questioned. But imagine our dismay when we adopt a practice faithfully and it turns out to be a dead end or to create more problems than it solves. What if a practice touted as best violates our "deepest convictions about the complexities of the writing process?"

Just as reflective teachers must question their own performance, we must be willing to question the methods accepted as best by the field of writing methods, an idea that may strike us as sacrilege. The very words *best practice* are loaded; if we aren't following best practice, aren't we by extension following worst practice? In addition, the term drips with authority. We may imagine that the process of determining best practice is mysterious but vaguely scientific; we picture labs with student control groups and teacher researchers in white coats behind one-way mirrors taking copious notes. We may assume that methods destined for the best practice label must undergo rigorous testing akin to the FDA or ADA processes of approval.

Our mental association of the term *best practice* with a doctor's white lab coat and its aura of infallibility is no accident. In the preface to their book *Best Practice: New Standards for Teaching and Learning in America's Schools*, Steven Zemelman, Harvey Daniels, and Arthur Hyde (1998) explain the origins of the term in education. In the late 1980s and early 1990s these authors were concerned that the school reform movement in Chicago would ignore important "issues of teaching and learning." They received a grant from the Joyce Foundation to create a newspaper focused on instruction, a newspaper they named *Best Practice 1*.

Why did we adopt the term "Best Practice," first for our newspaper and now for this book? We borrowed the expression, of course, from the professions of medicine and law, where "good practice" and "best practice" are everyday phrases used to describe solid, reputable, state-of-the-art work in a field. If a practitioner is following best practice standards, he or she is aware of current research and consistently offers clients the full benefits of the latest knowledge, technology, and procedures. If a doctor, for example, does not follow contemporary standards and a case turns out badly, peers may criticize his or her decisions and treatments by saying something like, "that was simply not best practice." (vii–viii)

Since the starting teacher salary at my public school is less than two hundred dollars higher than Salary Wizard's estimation of the average starting salary of a light truck driver in a city near mine, I understand Zemelman, Hyde, and Daniels' urge to link our work with the work of doctors and lawyers. Perhaps this association could lend us a measure of the respect and pay we're surely due. But we

would be wise to examine our metaphors lest they lead us to a place we never intended to go. The authority conferred upon educational best practice by its association with the medical field has far-reaching implications. The No Child Left Behind Act calls for closing the achievement gap using “effective, scientifically based instructional strategies” (Public Law 107–110, Section 1001, (9)). Like drugs, if teaching strategies, methods, and assessments can be proven effective in clinical trials, don’t we want them administered properly and consistently? If a doctor can be sued for failing to order a CAT scan when indicated, shouldn’t teachers be held accountable for dispensing proven methods?

Best practice, then, becomes a supporting argument for mandating increasingly specific practices. Throughout New York City’s Public School Region 10 in 2005, English language arts (ELA) teachers were required to display books in crates cover side out. Literacy Coaches at Intermediate School (IS) 172 did sporadic checks, reprimanding the unruly teacher who might have had the audacity to store her books in a bookcase spine side out. It’s hard to argue with the good sense in displaying books in crates cover side out; besides being aesthetically pleasing, such an arrangement makes books more visible. And book crates are definitely more portable than bookcases; imagine all the possible exciting configurations! The problem comes when this arrangement is enforced. “I spend too much of my prep period arranging and rearranging my class library,” says my sister Judy, an ELA teacher at IS 172. “I’d rather spend my time reading young adult literature so I can match books with my students.” But forget the power of a well-timed book suggestion; conclusions from the studies on personal recommendation aren’t in yet.

If only good teaching were as easy as following an approved list of prescriptions. Unfortunately, children are not bacteria to be obliterated by the correct dose of penicillin, and classes are not control groups whose every variable can be isolated. Since human beings are complex and class dynamics often surprising, teachers must be free to explore a wide range of possible approaches. Imagine the difference to teachers at IS 172 if the term *best practice* had been changed to *promising practice*. While the word *best* assumes a fixed canon of methods and closes off the possibility of other ways, the word *promising* offers the possibility of exploration. The question would no longer be, “Are you following best practice?” but “Are you exploring, discovering, and creating practices with promise?”

Even if we accept the medical model of mandating proven methods, we cannot be bullied into accepting the scientific infallibility of best practice; the blind acceptance of best practice in medicine has sometimes ended in disaster. Consider Thalidomide. By the 1960s, doctors in Europe and Canada regularly prescribed Thalidomide to pregnant women to help them with morning sickness. Studies indicated that the drug, unlike barbiturates, could be administered without toxic side effects. Although the first reported “Thalidomide baby” was born with severely deformed ears in 1956, it was not until 1961 when a German doctor determined that 50 percent of children born with deformities had been exposed to Thalidomide in the first trimester of pregnancy that countries began to pull the drug from the market.

Certainly, no teaching method that currently enjoys the label of best practice will result in death or deformity, at least of the body. The example of Thalidomide as it illustrates the argument for skepticism of best practice in education is hyperbole. But its legacy should give teachers the courage to question best practice.

If it is true that the field of writing methods can lead us astray, and if it is true that we cannot blindly accept teaching methods touted as “best,” how do we know which daily practices to question? We certainly don’t have time to investigate every study published for every classroom practice we have ever tried. Perhaps, as Chris Anson (1989) suggests, we can begin by questioning practices that violate our “deepest convictions about the complexities of the writing process.”

However, in the middle of the school year, it is easy to forget to reflect on our deepest convictions about the complexities of the writing process. The school systems in which we teach do not consider reflection to be a part of our job. Besides study hall duty, lunch duty, school improvement duty, curriculum council duty, and our own personal duty to use the restroom at least once a week, our job is to usher over one hundred students in and out of our rooms every day, to prepare them for state writing tests, the AP test, the ACT, the SAT, tomorrow’s unit test, the real world, and adulthood. Without time to think and reflect on why we began teaching English in the first place, our daily practices are vulnerable to the demands of local, state, and national politics—demands that may have nothing to do with the reasons and ways that human beings connect powerfully with language. If we are to create the conditions in which our students can experience the rich power of language, we must constantly remind ourselves of our own deepest convictions about writing. Once we

remember these, identifying practices that violate these convictions is the easy part.

You have probably picked up this book because something about rubrics violates your deepest convictions about the complexities of the writing process. If so, your misgivings have likely had little validation. While leaders in the field of writing assessment such as Patricia Lynne, Bob Broad, and Brian Huot have done invaluable work in pushing writing assessment at the university level beyond the limits of rubrics, I've not encountered a single book or article intended for K through 12 writing teachers that critiques rubrics. Without validation, our misgivings fade into resignation. *Rethinking Rubrics* is an attempt to articulate and explore how rubrics may violate the complexities of the writing process so that we can begin our search for more promising practices.

My Troubles ¹ with Rubrics

*I collect rubrics. I love them. They are as dear to me
as beanie babies, barbie dolls, mugs, key chains,
NCTE memorabilia, and dust bunnies under my bed.*

—KIT GORRELL (1998)

My introduction to rubrics came in second grade when my swim teacher used a series of statements in nifty boxes to assess my front crawl and elementary backstroke. I remember proudly reading the phrases and numbers praising the grace of my strokes to my mother and pitying the poor children whose flutter kicks were still progressing, understanding even then that the low numbers attached to the these boxes belied their euphemistic phrasing. Rubrics followed me into teacher training and beyond. I still obsess over why my student teacher performance rubric rated my classroom management skills as “adequate” rather than “excellent.” Did I not use proximity and make eye contact with misbehaving students? Did I not flick the lights on and off as a last resort? When I became a new teacher in Michigan, I was quickly introduced to the four-point rubric used by the Michigan Educational Assessment Program (MEAP) to assess student writing, and was just as quickly introduced to the MEAP’s six-point rubric when it changed the following year. I once promised my students several bonus points if they would simply staple the process paper rubric I had distributed to their final how-to drafts. I have been asked during several interviews if I can effectively use rubrics. Not only can I effectively use rubrics, but I also cannot escape them.

Like any other idea in education that can be expressed in a spreadsheet, rubrics have been co-opted for profit; textbooks and canned instruction programs come complete with pre-made assessment rubrics. I typed www.rubric.com in the address box of my browser, and was pleased not only to learn everything I ever wanted to know

about the joys of rubrics, but I was also able to download The Rubricator™, a software program that would allow me to quickly create my own rubric, link performance tasks to any national standard, and choose between many pleasing layouts at the click of an icon. All for only \$29.95, last time I checked. At a recent National Council of Teachers of English (NCTE) conference, I was invited by a large sign in the exhibit hall to “Take the Rubric Grading Challenge.” While I failed the challenge and was not sent a free copy of the Rubric Converter™, I was urged to order one posthaste for one low price. Rubrics are lucrative in addition to being ubiquitous.

Rubrics’ position as the latest sacred cow of writing assessment is no accident; rubrics make powerful promises. They promise to save time. They promise to boil a messy process down to four to six rows of nice, neat, organized little boxes. Who can resist their wiles? They seduce us with their appearance of simplicity and objectivity and then secure their place in our repertoire of assessment techniques with their claim to help us clarify our goals and guide students through the difficult and complex task of writing.

Yet, if you’re anything like me, you have mixed feelings about rubrics. You’ve used them. In fact, sometimes you really like them. Still, you’ve picked up this book because something about rubrics violates your “deepest conviction about the complexities of the writing process” (Anson 1989) and you question your own use of rubrics: are they really all they’re cracked up to be?

My own dilemma with rubrics crystallized when I taught a writing course at my district’s alternative education school. Most students came to my class professing a profound dislike for writing. With lives complicated by varying combinations of poverty, pregnancy, drugs, probation, school failure, and refusal or inability to fit in, the safety of formulas and worksheets was strangely comforting; a complicated process such as writing left them confused and leery. But I didn’t believe that students actually learned much or felt good about the time they spent filling out word searches. Knowing how important creative nonfiction writing had been to me, I created a Personal Narratives class and personally recruited (begged, pleaded, bribed) students to join.

When Felicity joined my Personal Narratives class as an eleventh grader she wrote about her grandfather’s death. Felicity’s writing was meaningful to her. She liked it, and it was organized clearly, beginning with a description of the moment she found out about her grandfather’s death followed by her reflection on his life and connection to her. She used vivid details, relating a funny and illustrative story about

a family tradition her grandfather had begun involving fishing poles, a song, and pennies thrown in a lake. The flashback was clear, the details relevant, the mechanics relatively error free. Felicity enjoyed writing it, and I enjoyed reading it. I graded Felicity's paper against the following rubric, which would be used by the State of Michigan to score my eleventh graders' writing ability on the MEAP (2005) later that year. I'd used this rubric before, and I'd never disagreed with anything in it.

Holistic Scorepoint Descriptions

Here is an explanation of what readers think about as they score your writing.

- 6 Writing is exceptionally engaging, clear, and focused. Ideas and content are thoroughly developed with relevant details and examples where appropriate. Organization and connections between ideas are well controlled, moving the reader smoothly and naturally through the text. The writer shows a mature command of language, including precise word choice that results in a compelling piece of writing. Tight control over language use and mastery of writing conventions contribute to the effect of the response.
- 5 The writing is engaging, clear, and focused. Ideas and content are well developed with relevant details and examples where appropriate. Organization and connections between ideas are controlled, moving the reader through the text. The writer shows a command of language, including precise word choice. The language is well controlled, and occasional lapses in writing conventions are hardly noticeable.
- 4 The writing is generally clear and focused. Ideas and content are developed with relevant details and examples where appropriate, although there may be some unevenness. The response is generally coherent, and its organization is functional. The writer's command of language, including word choice, supports meaning. Lapses in writing conventions are not distracting.
- 3 The writing is somewhat clear and focused. Ideas and content are developed with limited or partially successful use of examples and details. There may be evidence of an organizational structure, but it may be artificial or ineffective. Incomplete mastery over writing conventions and language use may interfere with meaning some of the time. Vocabulary may be basic.

- 2 The writing is only occasionally clear and focused. Ideas and content are underdeveloped. There may be little evidence of organizational structure. Vocabulary may be limited. Limited control over writing conventions may make the writing difficult to understand.
- 1 The writing is generally unclear and unfocused. Ideas and content are not developed or connected. There may be no noticeable organizational structure. Lack of control over writing conventions may make the writing difficult to understand.

Scoring Felicity's paper was relatively straightforward. While I felt that her "command of language, including word choice, supports meaning," and placed her in the 4 category, I gave her a 6 for ideas, which were "... thoroughly developed with relevant details and examples where appropriate." I wasn't sure if I was supposed to average the scores, but since everything else fell within the 5 range, I felt justified in authoritatively circling the 5.

Krystal, on the other hand, had confessed when she handed in her paper that she wasn't even sure what she was writing about; her piece involved thunderstorms, a trip to Texas, and a few tidbits about being teased by her cousins. The writing was a bit of a mess: inconsistent paragraphing, full of unintended fragments, unclear transitions, and rife with spelling, punctuation, and sentence structure errors. I could tell that Krystal's paper was going to score badly on the MEAP rubric. But Krystal's paper moved me deeply, which I at first struggled to understand. I had read enough badly written student papers about traumatic events to know when my reaction to the disclosure was interfering with my assessment of the writing. But Krystal wasn't recalling a particularly traumatic event; her writing itself moved me.

I hesitated to circle the 2—a failing score as far as the State of Michigan was concerned. It sounded so harsh—*The writing is only occasionally clear and focused. Ideas and content are underdeveloped. There may be little evidence of organizational structure. Vocabulary may be limited. Limited control over writing conventions may make the writing difficult to understand.* But even if I bumped organization up to a 3, I could just as easily reduce the convention description to a 1—the grammar, spelling, and punctuation *did* make the writing difficult to understand, averaging out to an overall score of 2.

I fully appreciated Krystal's struggle with mechanics. Krystal did as well; she often approached me and asked me to help her with a word or sentence. I didn't want to dismiss the importance of mechanics to

Krystal's development as a writer; she needed to use words and structure sentences and paragraphs confidently and skillfully in order to better say what she thought. But if someone had asked me to hand over the best paper in the batch, I would have handed over Krystal's without hesitation. In fact, I had shared Krystal's writing with two colleagues and a friend earlier that day, telling them that *this* paper was what made being a writing teacher the best job in the world.

When a thunder storm comes it gets as cold as winter
 The clouds turn gray
 Winds blows as hard as a tornado
 The birds stop cherping all the animals seem to disappear
 As if their being hunted by a wild animal
 You haer the thunder echo from so far away
 Seams clam then drastic at the same time sprinkles hit the ground
 The rain comes down harder and harder
 You hear the rain making a soft noise then it gets louder . . .
 Takes me back to a farm in Texas some time in April around Easter
 Where all my familiee got together to celebrat . . .
 All I can remember is a flat surface of the horizon that seemed it
 never ended
 Beyond was the biggest sun I have ever saw and the warmth seemed
 so nice . . .
 I never liked the familiee get together things
 because my cousins never really made us feel wanted
 Never made any conversation with my brother sister and I
 Maybe because we talked different I don't know
 But it seems every were I go I always some how feel missed place
 So when I think of sorrow
 It reminds me of a thunder storms
 Because of the colors gray and darkenss
 Because of the noises are loud and hectic
 The rain is like some one crying for an answer
 It reminds me of everything bad or sad that has happened to me
 I am thankful for everything because in away when any one makes
 me feel missed place
 Or unwanted I don't let it get the beast of me because I feel like that
 has happened to me a lot when I was a kid and know times have
 to change for me and I need to stay strong
 So maybe that's why I think of my self as independent and I don't
 think of myself as a follower or a leader and know I am happy

for who I am and it seems like no matter what the situation is I always find away to keep my self conferrable . . .

After a thunder storm the rain stops the clouds are blue and the sun is shining so bright and the birds start chirping and all the rain has disappeared . . .

Even in its rough state, I found Krystal's writing more exciting than many polished personal narratives I'd received when I taught upper-level college prep writing classes at the traditional high school. But nothing in the MEAP rubric reflected my excitement about Krystal's paper.

Perhaps I was using the wrong rubric. Even though my students' writing (and my teaching) would eventually be judged by this rubric, I was not required by my department or district to use it in my classes. In the interview for my teaching position three years earlier, I'd been asked if I was familiar with the 6+1 Trait® rubric, which was developed by a group of teachers led by Vicki Spandel in the early 1980s. I'd never used this rubric, but I'd been part of the NCTE listserv for several years and had followed several discussions about it. I'd even looked it up the night before the interview, and impressed at least myself when I'd rattled off the 6+1 Traits®: voice, sentence fluency, presentation, conventions, ideas, word choice, and organization. I liked the focus on voice; perhaps it would help me reconcile how I felt about Krystal's writing with her score.

I found the rubric on the Northwest Regional Educational Laboratories (2005) website. While I was shocked that it was fourteen double-spaced pages long (weren't rubrics supposed to streamline the response process?), I printed it. At first, I was hopeful. I could see that Krystal's writing did meet many criteria for the highest score in voice, including, "The writer takes a *risk* by the inclusion of personal details that reveal the person behind the words," and "*Narrative* writing is personal and engaging, and makes you think about the author's ideas or point of view." Elsewhere, however, Krystal's paper scored worse than it had according to the MEAP rubric. For ideas, her paper earned the lowest score, because, "As yet, the paper has no clear sense of purpose or central theme. To extract meaning from the text, the reader must make inferences based on sketchy or missing details." For organization, a score of 1 applied again, since, "The writing lacks a clear sense of direction. Ideas, details, or events seem strung together in a loose or random fashion; there is no identifiable internal structure."

Her performance on the sentence fluency and word choice traits fared no better; both scores were a 1 since “Sentences are *choppy, incomplete, rambling or awkward*; they need work. . . . There is little to *no “sentence sense”* present. . . . Problems with language *leave the reader wondering*. . . . Many of the *words just don’t work* in this piece. . . . *Language is used incorrectly* making the message secondary to the misfires with the words. . . . *Limited vocabulary and/or misused parts of speech* seriously impair understanding.”

But even with confirmation from two different rubrics, I couldn’t bring myself to fail Krystal’s paper. I wasn’t even tempted to fudge a little bit and slide her by with a D. Did I want to give her an A? I wasn’t sure, but I wanted to celebrate Krystal’s writing, to read it to the class and say, “Wasn’t that good?” to hang it on my wall and glance at it when I questioned whether I should stay in teaching or get that cosmetology license I’d always wanted. Mechanics aside, even the broad range of descriptive statements that the 6 +1 Trait® rubric provided didn’t capture the essence of my reaction to Krystal’s writing.

I reread the MEAP rubric to make sure I wasn’t missing something; how could my reactions and the rubric be so out of sync? But, apparently, my reactions didn’t matter much to the MEAP. As the first line reminded me, “Here is an explanation of what readers think about as they score your writing.” Obviously, I hadn’t thought about these things when I read Krystal’s paper, and I began to feel that this seemingly simple explanation for the student was really a mandate and a warning to the scorer—*No matter what you really think, you will think about these things as you read and score this paper*. Well, I’m not overly fond of a mandate, so I put the rubrics aside to figure out my response as I read Krystal’s paper.

I knew that Krystal’s associative leaps—thunderstorms to Texas to language barriers to feeling misplaced—would confuse some readers, but I loved how her words hinted at something more. I could imagine Krystal’s piece as poetry, and if I encouraged her to turn her writing into poetry, I wouldn’t have to mark her down as much for sentence structure problems. But I thought that the writing worked well as a narrative. I loved the poetic quality of Sandra Cisneros’ and Annie Dillard’s prose, and I didn’t think that associative writing and experimentation with sentence structure should be banned from narrative and relegated to poetry. While I imagined that some of the revision process for Krystal might involve elaborating, I also thought that part of the power of her piece hinged on its loose suggestive quality. Despite both rubrics’ focus on explicit, easy transitions, I hoped that

she chose to preserve much of the unconscious associative nature of her writing.

I also valued the exploration of context and marginalization I saw in Krystal's paper. While she didn't say it explicitly, Krystal's mention of feeling misplaced, difficulty with accents, the trip to Texas, and her connection to storms and the sun suggested her family's history as migrant workers. I also saw her many writing errors in the context of her displacement; she'd spent several months of many school years visiting her father in Ohio, so she'd not spent many complete grades at the same school. I'd spent a summer teaching English as a Second Language on an air-conditioned school bus with benches replaced by desks. We parked this roaming classroom at different migrant camps in the area every evening, and men and women who had just worked grueling fourteen-hour days picking carrots and asparagus boarded our makeshift school every evening to learn English. Some of these men and women had been professionals in Mexico; one had been an architect, another a lawyer, yet they earned more on Michigan's farms. Many of them told me that they worried about their children. Since they traveled from Michigan in the summers to Texas or Florida every winter to make ends meet, they wondered how their children would keep up in school. Krystal's personal story mirrored a larger societal context that I found important and intriguing. Krystal's story was not only important for her to write, but it was also important for us to read.

My appreciation of marginalization and the unconscious associative qualities I saw in Krystal's paper reminded me of my values as a reader of literature. I was fascinated by the texts and discussion I found in my postmodern literature courses in college. I fell in love with *Coming Through Slaughter* by Michael Ondaatje, a beautiful and unusual work of historical fiction based on the life of the jazz trumpeter, Buddy Bolden. At one point, Ondaatje injects a brief description of a woman cutting carrots. She cuts rapidly, caught in the repetitive movement of knife on carrot. However, the moment she thinks what she is doing, she loses control and cuts her finger. This passage has nothing to do with the plot of the story; the woman doesn't show up in any other section. I loved the initial confusion I had when I read this section; its lyricism kept my attention and I kept trying to figure out what it had to do with the story. My subsequent realization that this passage connected to a broader theme in the book—the allure and consequences of self-consciousness—felt like a victory. I'd solved some kind of puzzle as a reader, and felt proud of myself.

In addition, I valued the fact that Krystal was thinking through her writing. Felicity's paper was meaningful and interesting, but she didn't discover anything new by writing it. She knew before she started writing that she loved her grandfather and missed him, understandings that she put into words quite well. But she didn't surprise herself, or her reader, as she wrote. Krystal, on the other hand, stumbled onto some rather large insights as she wrote. While her description of thunderstorms was interesting in itself, she surprised herself and me as she began to connect the sound of the rain to her search for an answer and reliance on herself in the midst of uncertainty and rejection. Her writing brought both of us somewhere new. While Felicity's paper was clean and solid, I didn't see anywhere else for her to go with it. Krystal's paper was loaded with potential. I admired her risk and wanted to encourage and affirm it.

The MEAP and 6+1 Trait® rubrics failed to recognize my values as a reader and Krystal's strengths as a writer. If my assessment prompted Krystal to revise, the categories of the rubric would have suggested that she organize her paper in a way that would have changed the loose, poetic structure of what she had begun to do. I remembered how vehemently some of my classmates in college had hated postmodern literature; the leaps and associations confused, annoyed, or offended them. Apparently these same classmates had gone on to write rubrics. But if I created my own rubric to include categories such as, "Loose, associative leaps," "Explores issues of marginalization," or "Potential," I would have been imposing my readerly preference on Felicity's paper. Felicity's way of being in the world and thought patterns are much different than mine or Krystal's, and asking her to write in a postmodern style wouldn't fit the experiences and ideas she wanted to express.

What was I to do? The problem must be me—I was tempted to take a Valium and think as MEAP instructed me to think, circling a score of 2 with the authority of the State of Michigan behind me. Who was I to contradict the rubric? After all, I'd spent a lifetime immersed in, assessed by, and preparing to use rubrics. But despite all of this, here I was, stuck on Krystal's paper, questioning whether rubrics reflected what I knew about the complexities of the writing and responding process.