The Trouble with Rubrics (#)
The Trouble with Rubrics
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 portfolio 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 might be shown as better: a horde of rubrics, or a retreat to grades, or a return to case studies, or nothing at all. Experienced teachers know that not all alternative approaches to authentic assessment are genuine. My growth came in realizing that my own 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 myxir of dissent I heard from thoughtful educators 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 students.”[1] To which the only appropriate response is: oh-oh.

First of all, something that’s commended to teachers as a handy strategy of self-justification during parent conferences [Look at all the houses! My, Grandmaw! How I could have given a rubric that anything but a B!] doesn’t seem particularly promising for inviting teachers to improve their practices, let alone rethink their premises. Second, I’ve been looking into an alternative to grading and authentic assessment from the point of view not of students but of the professors who dare to present that what they’re doing is 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 very different way of judging people, finally delivered to language arts. 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 as long as you’re willing to gut the curriculum and turn the school into a test-preparation factory, so it’s possible to get a bunch of people to agree on what rating to give an assignment as long as 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 is objectionable in its own right. But it’s also harmful in a very practical sense. In an important article published in 1999, Linda Nathan, 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 scores is ever more easily achieved with regard to such matters as spelling and organization,” these are the characteristics that will likely find favor in a robotized classroom. Nathan cites research showing that “compliance with the rubric tended to yield higher scores but produced ‘seamless’ writing.”[2]

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 in which teachers benefit from consulting a rubric briefly 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 use them. The phenomenon I quoted earlier, who boasted of efficient scoring and convenient self-assessment, also wants us to employ these guides so that students will know ahead of time exactly how their projects will be evaluated. In the absence of this proposition, a girl who didn’t like rubrics is quoted as complaining, “If you get something wrong, your teacher can prove you wrong what you were supposed to do.”[3] Here we’re invited to have a good laugh at the 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 encourages me to view the quality of teachers’ work as secondary, 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.” She adds, “These days, when I tell them what to do,” she adds, “They don’t have confidence in their thinking or writing skills and seem unwilling to really take risks.”[4]

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+, or more likely a zero, is an end product. Only those who have to actually do the teaching are likely to question whether the whole enterprise has served. This is a theme sounded by Maja Wilson in her extraordinary book, The Schools Our Children Deserve (Houghton Mifflin, 1993) and Punished by Rewards (Houghton Mifflin, 1984), as well as in “From Boring to De-Grading,” High School Magazine, March 2006. (Continued on next page)

The fatal flaw in this logic 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 you’re reading for (example, to puzzle out why a character made a certain decision), and thinking about your own proficiency at reading. “Only extraordinary education is concerned with learning,” the writer Marilyn French once observed, “while most is concerned with achieving: and for young minds, these two are very nearly opposites.”[5] 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. Students even then that too much attention 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 the outcome to innate ability and other factors thought to be beyond one’s control.[7] 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 doing. I’m thinking about what I should be doing.” In fact, I’m not sure what to make of this conclusion. I’m not sure what to make of this conclusion. I’m not sure what to make of this conclusion. I’m not sure what to make of this conclusion. I’m not sure what to make of this conclusion. I’m not sure what to make of this conclusion. I’m not sure what to make of this conclusion.

In any case, 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. 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+, or more likely a zero, is an end product. Only those who have to actually do the teaching are likely to question whether the whole enterprise has served.

What all this means is that improving the design of rubrics, or inventing our own, won’t solve the problem because the problem is inherent to the very idea of rubrics and the goals they serve. This is a theme sounded by Maja Wilson in her extraordinary book, Punished by Rewards (Houghton Mifflin, 1984), as well as in “From Boring to De-Grading,” High School Magazine, March 2006. (Continued on next page)

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 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).

Neither nor our assessment strategies can be simultaneously devoted to helping all students improve and to sorting them into winners and losers. That’s why we have to do more than reconsider rubrics. We have to reassess the whole enterprise of assessment, the goal being to make sure it’s consistent with the reason we decided to go into teaching in the first place.

NOTES
2. I review this research in Punished by Rewards (Houghton Mifflin, 1984) and The Schools for our Children: (Houghton Mifflin, 1989), as well as in “From Boring to De-Grading,” High School Magazine, March 2006.

Another educator cited this same observation and adds: “Rubbish enough to give rubrics a closer look!” It’s also quoted on the Rubistar website, which is a sort of on-line rubric-o-matic.
5. Mindy Nathan, personal communication, October 26, 2006. 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 “competence.” As one 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 crush a semester of shared effort 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 are the greatest value to me.”

7. For more on the distinction between performance and learning – and the detrimental effects of an excessive focus on performance – see The Schools for our Children, chap. 2, which reviews research by Carol Dweck, Carole Ames, Carol Midgley, John Nicholls, and others.
8. Quoted in Natalie Perscheid and Carolyn Coovert, “Growing Beyond Grades,” Educational Leadership, October 2004, pp. 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 oversimplified – seem to have missed the former one.

Maja Wilson, Punished by Rewards (Assessment, 1984).