Paper presented at the ARTIST Roundtable Conference on Assessment in Statistics
held at Lawrence University, August 1-4, 2004

AGAINST FAIRNESS

ARTIST Roundtable Conference on Assessment Appleton, WI August 1, 2004

I'm supposed to say it's a pleasure to be here, and truly it is. But also, at the same time, I'm really disappointed. From the program, this looks like an unusual and wonderful conference, and I'm going to miss out. When I first accepted Beth's invitation to speak here, I had in mind the pleasant thought that of course I would stay for the whole conference and get to hear talks from a lot of people whose ideas I like and whose work I admire. Unfortunately, over the intervening months since then, unforeseen events have taken that pleasant idea and shoved it down the garbage disposal of harsh reality. First, I found that a standing summer teaching responsibility had been rescheduled. As a result, I spent all last week at Ohio State, and got back to Massachusetts only night before last. I'd also found out that I'll have to be leaving for Toronto this Wednesday in order to meet my obligations there. The disappointing upshot is that in order to have a day at home with my family, I'm flying back to Massachusetts tomorrow. This truncation of my time here is not only disappointing to me, but it also turns me into a kind of after dinner panda: an animal that eats, talks, and leaves.

Having just confessed that I can't be here to listen to you, I'm now going to give you four reasons why you should retaliate by not listening to me. Slouch back in your chair, and let your mind wander. Think your own thoughts. Ignore me.

Here are the four reasons: First, *ignorance*. What I'm going to talk about is unsupported by any foundation in research, untainted by any knowledge of the assessment literature, and bloated with undocumented opinion. Second: *arrogance*. I feel totally free of any obligation to be acknowledge opposing views, and I intend to exploit that sense of irresponsibility shamelessly. Third: *hypocrisy*. I'm going to talk about an approach to assessment that I'm too cowardly or too unimaginative or too lazy to try next semester, because next fall I'll be facing classes of 40, instead of the smaller classes of a dozen or so that I usually teach. Finally, *overreaching*: You assessment folks are a tough crowd, in that you are by inclination innovative, even iconoclastic, so it is really hard to say something you haven't already thought of on your own. I'm sure that some version of what I'm going to say is already familiar to many of you. All I can do is to try to be especially flagrant about overstating the case.

Well, there you have it: ignorance, arrogance, hypocrisy, and overreaching … Do you think I could get elected to high public office?

My remarks tonight address three concerns: fairness, grade inflation, and a third concern that for now I'll simply label "Roger." Each of the three concerns is linked to a corresponding attitude toward assessment. Taken together, the three concerns suggest an approach to assessment that I've come to use in my smaller, more advanced classes. I remind you, though, that so far, I've been too cowardly, unimaginative, and lazy to try the same approach in my larger classes.

1. Fairness.

My first and main concern tonight is fairness … I'm against it. I'm going to argue that a misdirected commitment to fairness is responsible for a variety of unintended evils in teaching and learning. If you read the Harry Potter books, think of fairness as the Dementor that oozes in and sucks all the joy out of learning.

I'm also going to argue that a misdirected notion of fairness is one of the main obstacles to authentic assessment. The bare bones argument is a simple one: Fairness involves comparing students with each other. Accepting fairness as a goal pushes us toward quantification, toward uniformity, and toward a view of learning as a zero-sum game.

If we quantify, it becomes easy, almost inevitable, to hold everyone to the same standard: your grade is based on your score. This nominally "fair" approach is easy to defend, in the sense that you can explain why one student's grade is higher or lower than another's, but it is not easy to defend in terms of its consequences for learning.

Think about two sets of consequences, for students, and for curriculum.

Students. First, the consequences for students. Students vary: their backgrounds vary; their interests vary; their learning goals vary; their learning styles vary.

Focus for a moment on the so-called "stronger" student. If everyone is graded on the same scale, teachers tend to feel less responsibility to challenge the better-prepared students. Students who start with more background or who pick things up quickly can expect to do well even if they work a lot less than students who start with less background or who learn more slowly. This is just how it is most of the time. "Stronger" students are misled into learning less than they might, and learning suffers. Our system penalizes talented students who should be learning to stretch themselves, and encourages a tendency to aim low. For these students, fairness encourages laziness and low standards for learning goals.

Now consider the so-called "weaker" students. If the assessment system is fair in the nominal sense, these students start with the deck already stacked against them. Quite likely they think from the beginning that it is impossible for them to do well, and they are resigned to a lower grade no matter what they do. They also may feel pressure to try to learn things they may not be ready for yet, or to settle for a superficial kind of learning, in order to keep from falling farther behind. Here, fairness encourages low self-esteem, and low standards for what it means to understand.

For all students, both the more prepared and the less prepared, both the quicker learners and the slower learners, misdirected notions of fairness encourage a sense of competition, discourage helping others, and encourage students to judge themselves and their accomplishments by comparing themselves with others, rather than judging themselves by what works best for them as individuals.

Curriculum. Let's turn now from students to curriculum. Because our concern with fairness encourages uniformity of expected outcomes, our system tends to foster an artificially compartmentalized curriculum. I find it useful to think about learning simultaneously in terms of the process and its content, so, as in quantum theory, learning has a kind of wave/particle duality. Fairness pushes things too far in the direction of the particles.

Consider the matter of prerequisites. Misdirected attention to equity says that we should measure all students using the same yardstick at the end of the semester. If everyone is to be graded on the same scale at the end, the system is fair only if we assume that everyone starts out at roughly the same place, which is what our grading system encourages us to expect of students. This means we're supposed to have a fairly strict

prerequisite structure. Everybody has to learn *abc* before the teacher is allowed to talk about *def*. You're supposed to take this course first, that course second.

Uniformity encourages us to deliver properly compartmentalized and sequenced quanta of knowledge; courses become standardized bags of curricular McNuggets; our curriculum becomes a succession of intellectual Happy Meals.

An alternative, wave, theory sees learning as a process. The best learning has a narrative flow, with a beginning, development, resolution, and integration. The learning occurs as our students weave the threads of new ideas into their existing understanding. Within reasonable limits, it shouldn't matter if students don't all start from the same place. Two different students taking the same course will inevitably get different things from it. We should embrace that inevitable difference, and try to see that each student gets as much as possible from our course, regardless of starting place.

Alternative. Instead of distorting assessment for the sake of fairness and uniformity, we can look to an alternative model in the form of the one-room school, with students in several different grades learning different things in parallel. Generations of teachers have made that model work. Why can't we think of each of our own classes as a more uniform version of the one-room school, with several levels together in a single room? The parallel is far from exact, of course, but as a metaphor for an alternative to what our system would have most college teachers do now, it can be useful. To make it work, we have to offer more options, for example, *more* homework, and more *kinds* of homework, than we expect any one student to do. Some will spend more time on drill-and-practice; others will be able to skip most of that, in order to spend more time on open-ended investigations and challenge problems. We'll need to encourage students to make their own choices from the work we offer, and we'll need to reassure them that no one is expected to do everything, and that we expect *them* to decide for themselves which things to skip and which things to do.

In short, then, we should *abandon fairness*. Instead, we should *assess students as individuals*, in the context of their individual backgrounds, interests, goals, and learning styles.

2. Grade inflation

The second of my three concerns is grade inflation ... I'm in favor of it. I'm going to argue that if we take authentic assessment seriously, then grade inflation is not just a likely consequence; rather, it is more: if our assessment is truly effective, then grade inflation is inevitable.

Learning, done right, is exciting, and gives students a sense of accomplishment. Our main goal as teachers is to help ensure this experience for our students, and a good approach to assessment should serve this goal.

What gets in the way? For many students, the obstacles are things like feeling bored, feeling competitive, feeling unfairly judged, feeling not up to the task, feeling the kind of time pressure that keeps you constantly in catch-up mode.

If we assess well, we will help our students to see clearly what their accomplishments are, and help them to enjoy the satisfaction of a clear-eyed recognition of their own progress, regardless of where they start from. Good assessment will help us to help them direct their efforts as effectively as possible, and they will accomplish more,

for the same effort. Their success will motivate them to work harder, and they will learn more as a result. If they are learning more, their grades *ought* to go up.

I used to think that I should judge *outcomes*, not effort. Students were wrong to feel that hard work alone was enough to entitle them to an *A*. If their final result wasn't up to the standards of the best in the class, their effort must not have been worthy enough.

I now give a lot more *A*s than I used to. I tell my students that if they and I can do a good job of directing their efforts, then how much they learn will depend mainly on how hard they work, and that everyone should be able to earn an *A* by putting in enough effort. In a typical class, although I try to be careful that students don't work so hard that they cheat their other courses of time, on evaluations almost all students describe their efforts in my course as either strenuous or fairly serious, with most saying strenuous, and the vast majority say that their enthusiasm for the subject has gone up. Students work harder, and enjoy it more, than they did back when I gave fewer *A*s.

So I urge you: *embrace grade inflation*. We needn't guarantee every student an *A*, and I don't think we should do that, but we *should* grade in a way that allows every student to regard an *A* as truly within reach, and we should assess in a way that allows both us and our students to feel that the *A*s we give have been truly earned, because our assessment has made both the learning and the connection to effort easy to see.

3. Roger

My third concern is the one I've called Roger. If you were at the last ICOTS in South Africa, and you were unable to avoid my after dinner talk there, then you already know about my *second* favorite cartoon. You folks here tonight are uniquely privileged: you get to hear about my *favorite* cartoon. It's by Gary Larson, and the caption is "Roger screws up." It's my favorite because it captures the essence of my inner life. Roger, the guy who screws up, is the percussionist in a symphony orchestra. In the cartoon, he stands poised behind three rows of string players, waiting to sound the triumphal crash of symbols that is supposed to provide the climax that the whole orchestra is building to. Acutely conscious of the weight of responsibility he bears, Roger has already raised his hands in readiness, and the balloon over his head shows him chanting to himself, "I won't screw up… I won't screw up …. I won't screw up." His facial expression is one of extreme concentration, and you can just imagine him counting the measures until his big moment, when, with all eyes upon him, he will bring his hands violently together. Although Roger doesn't know it yet, he is destined to blow it. A look at his hands reveals the disaster to come: Roger's left hand clasps a symbol, but poor Roger's right hand is empty. There will be no triumphal crash. Despite his concentrated determination, Roger will screw up.

Authenticity. What does Roger have to do with assessment? I'm going to answer my question with another quention: "Which is better, to be sincere, or to be authentic?" Casually, we might be inclined to think of "sincere" and "authentic" as roughly synonymous, but a look at their origins suggests otherwise. I'm going to argue that, in the original senses of these words, Roger is so preoccupied with being sincere that he has sacrificed his authenticity, and that, more generally, a misdirected sense of fairness in assessment encourages our students to make the same sort of unfortunate sacrifice.

At one time in the history of our language, to be sincere meant to be free from blemish, literally, "without stain." ("I won't screw up … I won't screw up.") Sincerity, in this older meaning, is fundamentally a conservative position, based on a minimax strategy: avoid blunders. Roger is so determined not to make a wrong noise at the wrong time that he ends up making no noise at all.

In his effort to be sincere, Roger has lost his authenticity. To be authentic is to be one's own agent, the author of one's deeds.

The tension between sincerity and authenticity parallels the tension between false fairness and effective assessment. Preoccupation with fairness leads to concern with grades, and ultimately, grades tend to be based mainly on "What did you get wrong?" *Fair* grading systems reward sincerity: being without stain. Authentic assessment is based on "What have you done well?" Authentic assessment should foster our students' sense of agency.

Hermit crabs Intellectually, we -- teachers and students both -- are all a bit like hermit crabs. In order to grow beyond our current capacities, we have to abandon the protective shell of attitudes, skills, and old habits of thought that we have become comfortable with, but that now confines us. Abandoning that shell leaves us temporarily vulnerable -- in particular, vulnerable to the embarrassment of screwing up -- as we develop a larger way of thinking. If we want to encourage our students to grow, we have to make it safe for them abandon their shells, safe for them to screw up. Our false notion of fairness works against creating the needed sense of safety. We substitute sincerity – being without stain – for authenticity --- being one's own agent.

Consequences. What does this mean in practice? Two things, I think. First, it means that our assessment should acknowledge learning as a process that involves feelings and people. One thing I do in this connection is ask students to write weekly journal entries in which they talk about their learning experience – their goals, satisfactions, frustrations, accomplishments, and so on.

Second, it means that our assessment should emphasize the positive. We should be sure to point out what a student has done well. Emphasizing the positive does *not* mean being a Pollyanna, however. It *does* mean that, rather than making an issue of everything a student has done wrong, we should for the most part focus attention on the few things we consider most important for the student to revise, and then make sure to provide an opportunity for revision.

I also avoid putting grades on written work. My premise is that if I can help a student direct her effort effectively, then how much she learns will depend on how much effort she puts in. In my comments, I try mainly to help her choose appropriate tasks and goals, I suggest revisions and extensions to the work she turns in, and I monitor effort, but in most cases I find that I don't have to comment on effort unless I think a student is working too hard.

Recap

To summarize, then, here is a short version, in three parts, of what I think I have learned about assessment, the three things that I regard as most important.

First, abandon fairness.
Try to assess a student's work in the context of that individual's own background, interests, goals, and learning styles.

Second, embrace grade inflation.
Emphasize effort and accomplishments. If good assessment leads to better learning, students deserve higher grades. Low grades are often mainly a symptom of the shortcomings of our system.

Third, surprise Roger.
Make it safe to screw up. Emphasize things done well and things that a student can fix, and provide the opportunity for revision. Value authenticity over sincerity.

To me, assessment is ultimately just a systematic way of paying attention to my students. This is not at all to say that if we just pay attention, assessment will take care of itself. The *systematic* part is essential, and that's where the science and craft come in. That's where we need research and experts. But systematic does *not* mean and *should* not mean fair. Fair, no; systematic, yes. Sincere, no; authentic, yes. In short, authentic assessment should assess authenticity.