Writing With Elbow

Belanoff, Pat

Published by Utah State University Press

Belanoff, Pat.  
Writing With Elbow.  
Project MUSE.  muse.jhu.edu/book/9312.

For additional information about this book
https://muse.jhu.edu/book/9312

For content related to this chapter
https://muse.jhu.edu/related_content?type=book&id=201984
So the most important point, then,
is that I am not arguing against judgment or evaluation.
I’m just arguing against that crude, simple way of representing judgment—distorting it, really—into a single number,
which means ranking people along a single continuum.

Peter Elbow

But if we drop the SAT,
by what means should we allot membership in the nation’s elite?

John Cloud

What really, as opposed to rhetorically,
transfixed late-twentieth-century America was
the precise calibration of a systematic national reward system,
which was what the testing and education regime had
become over half a century.

Nicholas Lemann

For Christmas last year, I asked for a copy of Nicholas Lemann’s recent book *The Big Test*. In that volume, Lemann, a staff writer for the *New Yorker*, details the founding and growth of ETS, the Educational Testing Service. Perhaps more than any single institution this century, ETS has shaped the very American culture of testing. In fact, one might argue that it’s not only shaped it, but also determined it—and therefore determined as well the education that tends to follow the test. Interestingly, as Lemann demonstrates, both ETS’s original purposes and its later developments were ideologically driven, often in surprising and explicit terms. Basically, the founders of ETS designed a program that in retrospect seems almost benign. They hoped to replace a system of advancement that in this country was based on inherited wealth and corporate and state connections with a simpler and—they believed—more equitable one driven by intellectual talent. Their vehicle for making change?

*Testing.*
Assessment is a large and technical area, and I’m not a professional. *

You have to wonder how it is that Peter Elbow got involved in an enterprise like writing assessment anyway. His graduate work is in Chaucer, much of his scholarly work explores topics like voice and freewriting, and as he says himself, he prefers to dissociate himself from claims made about his assessment expertise. In spite of these facts (or as we’ll see, perhaps because of them), Peter Elbow has become over the last fifteen years a leading figure in collegiate and university writing assessment—challenging the historic sorting function of assessment, advocating portfolios, and developing and advocating new ways to grade student work. In spite of reservations about writing assessment, he participated in the creation of the CCCC “Statement on Writing Assessment.” In spite of reservations about outcomes assessment, he is working with leaders of the WPA Outcomes Group, even if it is to help them consider reasons why we might not want a national statement of outcomes. His vehicle for making change?

Teaching.

Embracing [New?] Contraries

There are certain terms that don’t permit dialogue; it may be that testing and teaching comprise such a pair. Or: even when motivated by the same general good intentions, testing and teaching rely on fundamentally different understandings of human behavior. Writing assessment, which both redefines testing and locates it in the specific field of writing (Yancey, “Looking Back”), is likewise benign, intending to bring teaching and testing together, to make them congruent with each other, to open each to the possibility of accommodation between their agents: the testers, the teachers. As a teacher, as someone who practices assessment, I hope and work for such accommodation. But as I think about the different ways writing assessment is constructed by these parties, I wonder if such accommodation can be created, much less sustained. And as you’ll see, I conclude this chapter still wondering.

Besides, even if we can create and sustain such accommodations, we will encounter yet another problem. Even the assessments linked to teaching and learning—like the ones offered by Peter Elbow—can produce consequences that contradict their intents, can produce effects that are at odds with learning and teaching both.
Grading and evaluation are not bad in themselves, but they are bad in their effects when they monopolize the scene of teaching and learning.

Aware of this frequent disjunction between intent and effect, however, we can compensate, first, by identifying such distortions of intent and, second, by undertaking to learn about and correct them.

Is there a claim in this text, after all? I think so, but it’s a caution as much as a claim. Writing assessment has benefited from Peter Elbow’s work, to be sure. It is now understood through a new lens, located firmly within a new rhetorical situation, one defined by personal interaction, by connections, by willingness to learn—a situation familiar to teachers, new to testers. If this legacy is to survive, these attributes must locate both teaching and assessment. As Peter Elbow argues, they are intimately connected.

To put my larger point in key terms: I suggested above that testing was one, teaching another. But as I’ve written this chapter, taught classes and partnered with faculty at a new university, and worked with K-12 teachers from Virginia Beach to southern California, a third, definitive term has emerged, one that distinguishes teaching from testing, that brings perspective to this history and this interpretation of our work as faculty and Peter’s contribution to that work.

Trusting.

I hear more voice in these passages; something rich and useful and interesting is going on there; can you get more of that?

In 1992, I decided that I would edit a collection on voice in writing, which became the NCTE collection *Voices on Voice*. Although I didn’t know Peter, I’d written him a letter inviting him to participate, he’d said yes, and he’d been helping me think about the form the collection might take and the kinds of offerings—like an annotated bibliography—that might make the volume valuable to readers. As part of developing the proposal for the book, I again wrote Peter, this time to ask him to review my draft of the introductory essay, although I use the term draft here—as they say—advisedly. What I sent him was a set of scribbled notes that wandered in a stream of consciousness mode around a topic I found intimidating. My central question was not, as one might expect, what will the reader encounter in this text?—but rather, by what authority am I doing this collection?

Foolishly, bravely, I sent the draft to Peter.

Quickly, helpfully, he replied, circling one idea, “I like this.”

It was an assessment moment.
Validity was the SAT’s weaker point, but nobody within the testing world questioned the test’s reliability.

An emphasis on testing, from an ETS perspective, was what assured fairness. In fact, testing assured more than fairness; it effected social justice. What the ETS founders—people like ETS President Henry Chauncey and Harvard President James Bryant Conant—saw prior to 1945 was a country that claimed democracy as a central value while practicing an elitism predicated on the advancement of the mandarin classes: those who were born into wealth, who then gathered at the Ivies (which functioned more as country clubs than as sites of intellectual inquiry), and who then “naturally” progressed into positions of leadership and influence. What Chauncey in particular saw—and Lemann’s argument in The Big Test is that the story of ETS is the story of Henry Chauncey—was that democracy would be better served by grooming leaders who had native talent and intelligence. How this process of meritocratic selection was congruent with democracy is an unanswered question for Chauncey, but what was clearly answered was the means by which leadership for a democracy would be created: through a system.

This is what Henry Chauncey wants to do in . . . postwar America: he wants to mount a vast scientific project that will categorize, sort, and route the entire population. It will be accomplished by administering a series of multiple-choice mental tests to everyone, and then by suggesting, on the basis of the scores, what each person’s role in society should be. . . . It will accomplish something not very different from what Chauncey’s Puritan ancestors came to the New World wanting to do—engender systematic moral grace in the place of wrong and disorder—but via twentieth century technical means. The vehicle through which he hopes to achieve all this is an aborning organization called the Educational Testing Service, purveyor of a test called the SAT. (Lemann 5)

According to Lemann, Chauncey’s quest embodied the quintessential and paradoxical American promise: parity for all, rewards for the quick and the smart. Chauncey’s contribution to the realization of the American dream, of course, was the system that would make that potential real, a system that engendered complete faith. Ironically, what’s as interesting is the absence of faith—in either democracy or education—that drives one toward a system.

Last year [2000] 44% of the kids who graduated from high school took it [the SAT], up from 41% in 1995. In all, more than 2 million students took the SAT in 2000. The second-biggest admissions test, the ACT, has 1.8 million takers. (Cloud [online])
What’s interesting as well is what happens to students when one’s faith is transferred from people to systems.

In postwar America, what happened was simple but profound: education changed. The role that the personal in educational access had historically played shrank, becoming replaced by normed information that assured fairness by its anonymity and its numbers: student as a GPA, a class rank, an SAT score, itself normed. Collectively, as a people, we became sorted, categorized, and ranked. To assure that such sorting was itself fair, we relied on science and on the apparatus it supplied. We talked not about what students were learning, or had learned, or might learn, but about four kinds of validity and about how to check for those; about ways to calculate reliability; about ways to translate raw scores into something meaningful—like percentile rankings that themselves provided another means of sorting. These systems promised efficiency, to be sure, but that was not their chief virtue: engineering society was that.

I see three distinct problems with ranking: it is inaccurate or unreliable; it gives no substantive feedback; and it is harmful to the atmosphere for teaching and learning.

In the process of engineering society, the ETS founders somehow lost a virtue that assessment might offer: its function of valuing what is.

An emphasis on testing almost invariably reflects a distrust of teachers.

In some ways, it may be that Peter is always telling the teaching story, and that assessment for him is necessarily simply another version of that story. If we think of a teaching story as narrative, then what Gregory Clark says about Burkanic narrative is precisely relevant here. Clark makes the point that for Burke, narrative is

not story so much as it is contextuality. Contextuality embeds ideas and identity alike in the particular and dynamic complexities that develop in relational life, and it denies any possibility of their anonymity or autonomy. (129)

The key terms here, I think, include denial [of] anonymity and autonomy: for Peter, assessment is always about context, about community, and about the personal. Social justice and fairness are linked, as they were for Conant and Chauncey, but the vehicle for effecting them is not scientific, quantitative, and systematic, but rhetorical.

Assessment, therefore, serves the ends of rhetoric.
Let’s do as little ranking and grading as we can. They are never fair and they undermine teaching and learning.

I said earlier that the terms testing and teaching seem to talk past each other; and it is the latter that underlies an Elbovian view of assessment. Listen to some of Peter’s most famous lines, particularly as they counterpoint the assessment-as-system engineered by ETS:

- Imagine the absurdity of trying to score a person with a single number. (Black et al. 53)
- In fact, portfolios may now finally give us the leverage we have needed to dislodge our overreliance on holistic scoring in general: our habit of using single numbers to rank complex performances along a single dimension. (“Will the Virtues” 46)
- In the end, then, I conclude that the least interesting questions we can ask of any text—by students or published authors—are questions of quality or evaluation. The most intellectually interesting work comes from asking and answering many of our most common analytic and academic questions—questions that invite us (though they do not require us) to step outside the mentality of evaluation. (“Taking Time Out” 17)
- Surely, most of us have learned that we don’t so much help people improve as persons by giving them constant diagnosis of strengths and weaknesses. We help them by engaging with them in serious and felt relationship. (“Will the Virtues” 54)

When Peter’s contribution to writing assessment is summarized, what appears obvious is less the vehicle most closely associated with his name—portfolio—and more the new rhetorical situation he seems to be defining here. In the classic assessment situation, a student, a test-taker, is assumed to be a given, to produce that which can be measured scientifically, and then, on the basis of that product, to be sorted. It’s a fixed set, a static situation. In contrast to this, Peter’s sense of assessment evokes a different rhetorical situation: one that is fluid, emerging, and personal. It’s a situation embodied in a felt relationship, a situation whose primary purpose is to help students learn. If students aren’t being helped, then assessment is superfluous, as Peter reminds us:

This brings up a metaphor or parable that always returns to mind. In my last year of college, I had an old beat up car; it worked but not well. Then, a few years later; same kind of thing. Then my third car came a
number of years later when I first had a full time teaching job. It was a VW bug in pretty good shape.

When I first got it and discovered it didn’t have a radiator, I was amazed and then gradually got a feeling of having been cheated. All my troubles with my two earlier cars had been with the radiator—and here, suddenly, I was discovering that a radiator wasn’t necessary for a car. It didn’t help make it go backwards or forwards. It was just something that most designs built in as if essential. That’s my parable of assessment. It’s not part of teaching—even though we are lulled into assuming it is. It doesn’t really help make people learn. (Elbow and Yancey 105)

Several points here are worth noting. First, the priority belongs to teaching, not to assessment. If assessment cannot aid in the teaching enterprise, then its value is limited. Second, Peter draws his evidence not from sophisticated scientific theories or complicated theoretical applications, but from everyday experience, in this case from the experience of owning a VW. Third, the experience is conveyed in the form of a parable, that is, in the form of a narrative that pretends to truth, a literary form familiar to English teachers.

Finally, Peter is quite clear about the fundamental purpose of assessment: to help students learn.

In only twenty years or so, we have twice changed the world of assessment, and we did it by resisting conventional practices of the testing community and setting an example of sound practices.

Peter is often described as the academic who is overly personal, who doesn’t understand politics, who focuses overmuch on the individual. Perhaps unwittingly, these critiques have always sounded an ironic note to me. On the one hand, of course, there’s the ring of truth about them: Peter’s gaze does seem to cast on the individual rather consistently. On the other hand, the idea that such focus isn’t itself another kind of political act is itself surprisingly naive.

The critique of Peter for his allegedly apolitical rhetorical stance fails to consider both some of the tenets of reformers at the same time it fails to understand the history of writing assessment in the last thirty years. The academics making the critique, for instance, are typically the first to observe that the personal is always political; is this not also so in this case? Likewise, and more telling, these academics have not joined Peter in thinking about, talking about, and enacting social change in the one arena where all change finally is estimated: that is, in the rhetorical situation of assessment. And I have to say—have to jump out of this text at this precise moment, much as a Victorian narrator to say—how completely baffled I am by this reluctance. In
my view, we can incorporate new pedagogies like collaborative learning, and we can introduce new technologies like computers, and we can specifically enact reform curricula like service learning, but *if we do not provide for these changes to be valued in appropriate terms, they cannot effect the promised reform*. Instead, the new is held hostage to the old, and reform is defeated. What Peter has done, then, in speaking a kind of truth to assessment power is to *begin changing the very terms* by which learning and teaching are valued—and he’s taken this up not as an expert, but as a teacher. Why?

> [W]e nonprofessionals can and should work on it because professionals have not reached definitive conclusions about the problem of how to assess writing (or anything else, I’d say). Also, decisions about assessment are often made by people even less professional than we, namely legislators.

From one view, most of the major crises in composition studies in the last half of the twentieth century—from the continuing disputes about CUNY admissions to the banishment of remediation in the California State System—have hinged on assessment. Put differently, assessment has been used as the (political) vehicle to exclude and even preclude certain students, and more particularly, certain kinds of students. If we do not engage in these terms, we are powerless to influence even the potential conditions that govern our students’ academic lives. Speaking from the rhetorical situation of his own experience and his own classroom, Peter assumes that he has power, even in assessment matters, and he seeks to use it judiciously.

Not because he is especially interested in assessment, but rather, because he’s interested in teaching.

> Portfolios have kicked back at testing itself—helping people rethink some central assumptions and practices.

Should the Elbovian legacy survive, particularly the legacy represented in portfolios, two major problems must be addressed: first, what I’ll call, borrowing from Catharine Lucas, distortion of effect, for both students and teachers; second, a kind of naïveté about how what we teachers do will be used, will in fact be systematized by a culture that worships ranking. Or: is it ever possible to get outside the system? Alternatively, could we work both outside and inside the system?

One difficulty with any kind of change that people are drawn to is that no change is directly replicable. Rather, people learn about something new—in the case of teachers, they learn about new ways of teaching like collaborative
learning, or new ways of assessing like portfolios—and they interpret the new within their own ways of being, and then apply the new just so, using the new to suit their needs, their sense of their students’ needs.

It’s a messy, asystematic process, this kind of change. At the same time, the new application can be fundamentally at odds with the design of the original. The controversies surrounding whole language illustrate this principle fairly well. In general, whole language advocates, like Yetta Goodman, don’t advise getting rid of phonemic or context clues altogether, although they do advocate inviting students to read real and whole texts. But my son’s first grade teacher—a woman devoted to whole language, she said—wouldn’t allow a phonics lesson or a spelling rule, with the predictable result that my son spelled worse at the end of the year than he did at the beginning. Was I in favor of abolishing whole language? No. Was I in favor of its being judiciously applied? Yes.

It’s the same story with portfolios, I’m afraid. The intent, as Peter says, is to promote both better teaching and learning, not to create a new maze to puzzle (or defeat) students—and not to create yet another exercise for them to complete as mindlessly as possible. And yet . . . there’s anecdotal evidence, at least, that sometimes, perhaps more often than we’d like, portfolio-as-exercise is what gets implemented. Elizabeth Metzger and Lizbeth Bryant report a student, for example, who claims to have beaten a portfolio system privileging revision, more specifically to “have botched a paper so it looked like I revised.” (7) Inviting students to underachieve (or worse, to misrepresent their talents) was not the intent of the portfolio, but it can be the effect.

Question: given that our intents and our effects don’t always match, what might we do to prevent such misapplications?

A second issue concerns the value students assign to portfolios even when the implementation “works.” Liz Spalding and Gail Cummins’s study of first-year students at the University of Kentucky points us toward some answers. The students in the study had all completed Kentucky’s compulsory twelfth-grade portfolio; what Spalding and Cummins wanted to know was how students understood both the portfolio and the processes that contributed to its composition—and regardless of how you parse them, the results discourage. For instance, “some two-thirds of the students stated that compiling the portfolio was not a useful activity.” (191) As disturbing were the comments that students articulated:
I actually only had two pieces that could be used in my portfolio when time neared to turn them in. So the week before, I wrote three pieces off the top of my head just to turn something in. I don’t consider myself a good writer, but I did get a “proficient” [score]. I feel everything was a waste of time. (182)

Another commented,

I think it was helpful to an extent, but there was just too much emphasis placed on it. Many times we took time out of class to do and discuss portfolio pieces. This took away from valuable class time and while we should’ve been learning something to help further our education, we were discussing how to make a better portfolio (184).

The concern suggested in these comments is not that they emerge from a mandated program. If that were the problem, then we could work toward eliminating, or at least minimizing, the effects of a forced portfolio. I wish, in fact, that the problem could be solved that easily, as un-easy as it may sound.

I think, instead, that the problem here is multiple.

One problem: the curriculum doesn’t always support the portfolio. If a colleague wants to introduce portfolios, do you require that a curriculum be in place first? Let’s reverse the question: doesn’t implementing portfolios and then reviewing them provide a collective and textured way to learn about curriculum and to talk about it? At the same time, the student may have a point: while we are in the process of learning—which itself is part of both teaching and assessing—will students find that the gap between intent and effect, between curriculum and assessment produces nothing more than a waste of time? Is there a way to avoid this problem?

Another problem: many students want the very education that portfolios are attempting to replace. They don’t want discussions of portfolio pieces; they don’t want their valuable class time wasted this way. For them, class time is valuable only when the teacher talks, when the teacher directs them, when the teacher identifies knowledge, when lines are clear and ambiguity is erased. Given this situation, what can we do to help students understand not only the contents of portfolios, but also their design and their subtexts?

A final problem, not unrelated: what’s true of students can also be true of teachers. Just as students cannot be simply given portfolios as guaranteed vehicle-of-learning, neither can teachers be given them as guaranteed vehicle-of-teaching. Wendy Bishop makes this point in her discussion of using portfolios with new TAs. One of them reflects on whether he’ll use them again—once he can control what he does in his own classroom:

I don’t know. I don’t think that the Portfolio thing is a total disaster. It just didn’t do anything for me this time around. If I wanted it to work for me, I guess I would have to re-structure my entire class plan. (“Going” 225)
Like the student above, this teacher finds that portfolios call for a re-structuring, a new way of understanding teaching and learning. Is there a way to help the teacher begin the re-structuring? What would motivate such a teacher? And for all teachers: isn’t re-structuring a dynamic fundamental to the teaching enterprise? Or, is it rather re-structuring as dynamic fundamental to the learning enterprise?

(“Ten,” mutter the guys when they see a pretty woman.)

Still, I want to argue, as I did in a recent CCC article, that speaking in the language of the assessment experts can also be useful. In other words, if we know that language, we can use numbers strategically. It’s not either/or: we need to model best practices, yes: that’s one means of accomplishing change. But perhaps we need as well, now and again, to link our work to the numbers the culture loves. To illustrate this, I want to cite some work that is currently taking place in Virginia Beach. Chris Jennings, a faculty member at Tidewater Community College and a former high school teacher in Virginia Beach City Schools, is directing a FIPSE project that we might want to consider as a model of both/and: fostering better teaching and learning, and using numbers as one kind of confirmational evidence. Here’s a basic outline of problem and new practice and results.

The problem: students from Virginia Beach high schools enroll in very high numbers at Tidewater Community College; the writing courses there fall into three categories: 003; 001; and 101. English 101 is where we’d want all students placed, since it’s the entry-level course; typically, 67% of the students enter into 003 or 001. Seems to be a problem, the folks there said.

The new practice: Teachers in a specific Virginia Beach high school, Salem High School, learn about and begin to use portfolios. The student population excludes AP students: it’s a general track. The teachers aren’t writers, nor do they practice a rich curriculum. But they volunteer to work with portfolios, and like dominoes falling one to the next, the teachers develop a writing process approach with their students; they begin to use reflection; they respond to writing differently. Useful as a defining concept, the portfolios become an afterthought. At the end of a year, portfolios are collected and scored according to a scoring guide created collaboratively by teachers at both Salem and Tidewater.

Assessment . . . defines our work to outsiders and . . . to ourselves and locates us within the faculty as well as within the larger culture.
(White 307)
The preliminary results: The numbers flip: 67% of the students are placed into 101. These same students score better than their peers do on the Virginia “Standards of Learning” (SOL) state twelfth grade test. These students enroll in higher numbers in both 4-year and 2-year colleges. They stay in college longer than Tidewater students ordinarily do.

Could this effort have failed? Yes. Might it still? Yes. It’s preliminary, but it’s promising. The numbers above speak to the first year. The second year, different students, some new teachers, but the same results. This way of learning “works."

Is the effort being held hostage to these numbers—the percentage who place into non-basic first-year comp, the number who scored well on the Virginia state test, the number who go to college? No. (Not yet.) The SOLs, for instance, weren’t intended to be part of the research. But the students have to take the test, so we checked the scores—just to see. The scores were good, better than anyone would have predicted, suggesting that if the curriculum is rich, the students’ work will be as well. That’s okay. In other words, the numbers do tell a story, and we shouldn’t be afraid to see what that story is and how it compares with the story we think we are seeing. This is especially so when the numbers don’t drive learning, but are used to back up what we’ve seen as we’ve reviewed the portfolios: new curriculum, new students’ voices, new genres.

(Dare I say it? New selves.)

At the same time: do I like numbers? No, not much. You don’t work with portfolios—as Peter’s work suggests—because you’re a fan of numbers. But numbers we’re probably going to be living with for awhile yet. I’m not willing to tailor the curriculum to the test and its numbers, no, but if I have to tolerate the numbers anyway, I’m willing to use them to verify a curriculum that encourages and rewards and helps students learn. And part of being able to do this is knowing enough about numbers—being enough of a psychometrician as well as a rhetorician—to be able to use them to foster good.

In the best case scenario, you see, we can use them to show that they aren’t needed, after all.

It’s the mark of good writers to like their writing.

In thinking about the issues here, in reading Peter’s writings, in continuing to work in both teaching and assessment, I’m struck by some contraries, some oppositions, some assumptions. I’m reading the Greenville News, where a letter to the editor comments on the SAT and how it cannot be other than undemocratic. The letter’s point is that (1) high achievement by all students is a laudable goal and (2) the SAT as a measure is fundamentally unable to show this (much less reward it), given its intent of showing quite the reverse. In other words,
because the SAT yields norm-referenced scores, “high-achievement” means above-average performance on the test. The letter, thus, rightly concludes that by definition “all students in all states could never realize [high achievement].”

This, then, is what Peter, like this reader, has understood all along. In teaching, we have a choice. We can teach students; if we do this, it’s possible that all students can achieve. We can sort students; if we do this, we apparently have two options. We can agree, on the one hand, to sum the numbers as Garrison Keillor does, so that all students are above average. We can understand, on the other hand, that we’ll find half of those students, regardless of context—background, development, classroom conditions, you name it—recorded as below average. More to the point, it’s highly unlikely that this recording will in any way alter their chances to learn in the future—except perhaps negatively.

Not least, in reviewing Peter’s writings on assessment, one word appears and re-appears.

Trust.

Something at odds with a system.

Because something that is personal, intuitive, human.

Something we bring to and take from our own teaching and learning.

Something we want to be for and to each other.

I’m observing an instructor this term as a part of a personnel procedure. It’s pretty formulaic: like many places, we don’t pay much, so the instructor is probably with us for the life of the contract. Still, I enjoy watching teachers, and I enjoy watching students, and in each observation I learn something.

Today, late November, it’s sunny, the students ready to put the term to rest. They are working on cases for a technical communication class; the instructor has brought in several samples. She passes them around. She asks the students to read them. She explains each one, carefully walking the students through first one case, then a second, finally a third. The cases: they aren’t difficult. The students like her personality: it’s engaging. But I observe as well: they aren’t engaged. Concluded, the class concludes—15 minutes early.

I’m walking out the door, wondering (as I always do when the full time isn’t used) what isn’t quite working here. I can’t seem to put my finger on it. The students seem interested. The teacher’s very articulate. The cases illustrated well. Suddenly: but the students, they weren’t trusted, were they? Why didn’t they read the cases? Think about them? Try to figure out what in them was illustrative?
Couldn't they have performed the readings?

Trust cuts in many directions.

Of course, all this—it’s not really about the SAT, you know? It’s about the relationship between and among testing, teaching, learning, and democracy. It is about the SAT as exemplar, and not quite as (single) villain. To see the SAT as villain is to fail to see that it merely represents what can go wrong, even when one’s intentions are worthy. It was a worthy thing to find a substitute for wealth as the ticket to success in America. The problem is that the alternative means—the SAT—hasn’t altered that fundamental reality. Because it’s designed to produce winners and losers, the SAT has simply produced substitutions, new winners and losers, most of whom (by the way) look remarkably like the originals. Likewise, when constructed as tests, good classroom practices can go equally awry: portfolios become less vehicles for learning, more a means of sorting and ranking of students and of surveillance of teachers. Not least, testing itself—regardless of type—continues unabated into every nook and cranny of education, with results that are predictably deleterious. The tests become the curriculum, especially as teacher salaries and student promotion become contingent upon them. Or perhaps it’s a variation on this theme of a nightmare connection between assessment and the classroom, as, increasingly, the curriculum is devoted to ways of taking these tests and passing them.

What we can do, Peter reminds us, is to do what we do well:

- help students learn to communicate
- work together to design assessments that capture what they do, not what they don’t do
- speak to our own experience, and, not least, trust

For example, in Chicago, the Consortium on Chicago School Research concluded that “Chicago’s regular year and summer school curricula were so closely geared to the Iowa test that it was impossible to distinguish real subject matter mastery for passing this particular test.” These findings are backed up by a recent poll in Texas which showed that only 27% of teachers in Texas felt that increased test scores reflected increased learning and higher quality teaching. 85% of teachers said that they neglected subjects not covered by the TAAS exam.

(Wellstone [online])
In short, portfolio assessment invites us to ask the real assessment questions: “What do we really want in successful students? What are we trying to produce?”

To say that a common teaching theme emerges in Peter’s assessment writing is to understate badly: Peter’s interest in assessment both begins and ends in the classroom. This point, I think, is not sufficiently understood. It’s not only that teaching *per se* interests Peter, even after all these years; it’s also that he understands teaching both to provide a common defense against testing and to permit a way to change testing. It’s also that he plotted this change *largely* against a culture invested in testing, simply by returning again and again—and yet one more time—to what we know best: *teaching*. In other words, he identified a set of incongruences—between what we say we want, like achievement, and our own practices, and he spoke about those as a teacher. This approach provided one way of reforming writing assessment in America.

My focus is on pedagogy and practice. My approach is not methodologically sophisticated; I am simply trying to think through my own evaluative practices.

What’s also not so well understood is that in his teaching concerns, Peter not only shares, but also anticipates the concerns of assessment specialists. They talk in the language of consequential validity as they focus on the link between assessment and curriculum; he talks about the importance of helping students improve. They mean the same.

[T]he least interesting and useful question to ask about any piece of writing is how good it is.

To observe this similarity, however, is not to suggest that they share much beyond this singular concern for effect. Motivated by divergent values, they call upon different apparati to enact their agendas. As important, despite/because of his outsider status, Peter’s agenda has changed assessment practice: Foucault can mean multiply, it seems. Still, precisely because Peter’s changes inform our teaching, it behooves all of us to do what Peter’s assessment work has reiterated: trust ourselves, trust our students, and speak to that trust.

**NOTE**

* The quotes from Peter Elbow are indicated by a consistent font style and size, and they appear without page references, for two reasons. First, many of the quotes
are layered into the text, and including citations will disrupt the tenor of the text, so it’s in part a move to preserve the alternate feeling of the chapter. Second, and as important, since I’m including quotes from across the Elbow canon, using them without citation emphasizes the fullness of the canon and the gestalt of the work.