

The Loneliness of the Long-Distance Test Scorer

DAN DIMAGGIO

Standardized testing has become central to education policy in the United States. After dramatically expanding in the wake of the No Child Left Behind Act, testing has been further enshrined by the Obama administration's \$3.4 billion "Race to the Top" grants. Given the ongoing debate over these policies, it might be useful to hear about the experiences of a hidden sector of the education workforce: those of us who make our living scoring these tests. Our viewpoint is instructive, as it reveals the many contradictions and absurdities built into a test-scoring system run by for-profit companies and beholden to school administrators and government officials with a stake in producing inflated numbers. Our experiences also provide insight into how the testing mania is stunting the development of millions of young minds.

I recently spent four months working for two test-scoring companies, scoring tens of thousands of papers, while routinely clocking up to seventy hours a week. This was my third straight year doing this job. While the reality of life as a test scorer has recently been chronicled by Todd Farley in his book *Making the Grades: My Misadventures in the Standardized Testing Industry*, a scathing insider's account of his fourteen years in the industry, I want to tell my story to affirm that Farley's indictment is rooted in experiences common throughout the test-scoring world.¹

"Wait, someone scores standardized tests? I thought those were all done by machines." This is usually the first response I get when I tell people I've been eking out a living as a test-scoring temp. The companies responsible for scoring standardized tests have not yet figured out a way to electronically process the varied handwriting and creative flourishes of millions of third to twelfth graders. Nor, to my knowledge, have they begun to outsource this work to India. Instead, every year, the written-response portions of innumerable standardized tests given across the country are scored by human beings—tens of thousands of us, a veritable army of temporary workers.

I often wonder who students (or teachers and parents, for that matter) picture scoring their papers. When I was a student, I envisioned

DAN DIMAGGIO (dan.dimaggio@gmail.com) is an independent writer, temp worker, and member of Socialist Alternative in Minneapolis, MN.

my tests being graded by qualified teachers in another part of the country, who taught the grade level and subject corresponding to the tests. This idea, it turns out, is as much a fantasy as imagining all the tests are being scored by machines.

Test scoring is a huge business, dominated by a few multinational corporations, which arrange the work in order to extract maximum profit. I was shocked when I found out that Pearson, the first company I worked for, also owned the *Financial Times*, *The Economist*, Penguin Books, and leading textbook publisher Prentice Hall. The CEO of Pearson, Marjorie Scardino, ranked seventeenth on the *Forbes* list of the one hundred most powerful women in the world in 2007.

Test-scoring companies make their money by hiring a temporary workforce each spring, people willing to work for low wages (generally \$11 to \$13 an hour), no benefits, and no hope of long-term employment—not exactly the most attractive conditions for trained and licensed educators. So all it takes to become a test scorer is a bachelor's degree, a lack of a steady job, and a willingness to throw independent thinking out the window and follow the absurd and ever-changing guidelines set by the test-scoring companies. Some of us scorers are retired teachers, but most are former office workers, former security guards, or former holders of any of the diverse array of jobs previously done by the currently unemployed. When I began working in test scoring three years ago, my first “team leader” was qualified to supervise, not because of his credentials in the field of education, but because he had been a low-level manager at a local Target.

In the test-scoring centers in which I have worked, located in downtown St. Paul and a Minneapolis suburb, the workforce has been overwhelmingly white—upwards of 90 percent. Meanwhile, in many of the school districts for which these scores matter the most—where officials will determine whether schools will be shut down, or kids will be held back, or teachers fired—the vast majority are students of color. As of 2005, 80 percent of students in the nation's twenty largest school districts were youth of color. The idea that these cultural barriers do not matter, since we are supposed to be grading all students by the same standard, seems far-fetched, to say the least. Perhaps it would be better to outsource the jobs to India, where the cultural gap might, in some ways, be smaller.

Many test scorers have been doing this job for years—sometimes a decade or more. Yet these are the ultimate in temporary, seasonal jobs. The Human Resources people who interview and hire you are temps, as are most of the supervisors. In one test-scoring center, even the office

space and computers were leased temporarily. Whenever I complained about these things, some coworker would inevitably say, “Hey, it beats working at Subway or McDonald’s.”

True, but does it inspire confidence to know that, for the people scoring the tests at the center of this nation’s education policy, the alternative is working in fast food? Or to know that, because of our low wages and lack of benefits, many test scorers have to work two jobs—delivering newspapers in the morning, hustling off to cashier or waitress at night, or, if you’re me (and plenty of others like me) heading home to start a second shift of test scoring for another company?

Company communications with test-scoring employees often feel like they have been lifted from a Kafka novel. Scorers working from home almost never talk to an actual human being. Pearson sends all its communications to home scorers via e-mail, now supplemented by automated phone calls telling you to check your inbox. After the start of a project, even these e-mails cease, and scorers are forced to check the project homepage on their own initiative to find out any important changes. Remarkably, for a company entrusted with assessing students’ educational performance, messages from Pearson contain a disturbing number of misspellings, incorrect dates, typos, and missing information. Pearson’s online video orientation, for example, warns scorers that they may face “civil lawshits” from sexual harassment. Error-free communications are rare. I was considering whether this was a fair assessment, when I received a message from Pearson with the subject “Pearson Fall 2010.” The link in the e-mail took me to a survey to find out my availability—for the spring of 2011.

Communications at scoring centers are hardly better. For example, test-scoring jobs never have a guaranteed end date. If you ask a supervisor when a job is going to be completed, you will get a puzzling response that “we don’t know how many papers are in the system, so we can’t say when we’ll be done.” This response persists, even though it’s pretty easy to calculate how many fifth graders there are in Pennsylvania and how long it will take to grade their papers, given our scoring rate. If we are lucky, we get twenty-four-hours notice before being told that a project is about to end and we should seek other work. Two hours notice is more common. In general, scorers are given no information beyond what is absolutely necessary to do the job.

What is the work itself like? In test-scoring centers, dozens of scorers sit in rows, staring at computer screens where students’ papers appear (after the papers have undergone some mysterious scanning process). I imagine that most students think their papers are being graded as if they

are the most important thing in the world. Yet every day, each scorer is expected to read hundreds of papers. So for all the months of preparation and the dozens of hours of class time spent writing practice essays, a student's writing probably will be processed and scored in about a minute.

Scoring is particularly rushed when scorers are paid by piece-rate, as is the case when you are scoring from home, where a growing part of the industry's work is done. At 30 to 70 cents per paper, depending on the test, the incentive, especially for a home worker, is to score as quickly as possible in order to earn any money: at 30 cents per paper, you have to score forty papers an hour to make \$12 an hour, and test scoring requires a lot of mental breaks. Presumably, the score-from-home model is more profitable for testing companies than setting up an office, especially since it avoids the prospect of overtime pay, the bane of existence for companies operating on tight deadlines. But overtime pay is a gift from heaven for impoverished test scorers; on one project, I worked in an office for twenty-three days straight, including numerous nine-hour days operating on four to five hours sleep—such was my excitement about overtime.

Yet scoring from home also brings with it an entirely new level of alienation. You may work on a month-long project without ever speaking to another human being, never mind seeing the children who actually wrote the papers. If you do speak to another person, it's at your own expense, since calling the supervisors at the test-scoring center takes time, and might cut into the precious moments you spend scoring (especially when you have to wait fifteen minutes for someone to answer, as happens routinely on some projects).

The piece-rate system also leads to some sinister math; I have often wondered how much money I lose for every trip to the bathroom, and debated taking my laptop there with me. And since you are only guaranteed employment until the papers run out, you are in a race against all your phantom coworkers to score as many papers as you can, as fast as possible. This cannot be good for quality, but as long as the statistics match up and the project finishes on time, the companies are happy. I did receive some automated warnings from Pearson that I was scoring too fast, while simultaneously receiving messages on the Pearson website to the effect that, "We're way behind! Log in as many hours as you can and score as much as possible!"

No matter at what pace scorers work, however, tests are not always scored with the utmost attentiveness. The work is mind numbing, so scorers have to invent ways to entertain themselves. The most common method seems to be staring blankly at the wall or into space for minutes

at a time. But at work this year, I discovered that no one would notice if I just read news articles while scoring tests. So every night, while scoring from home, I would surf the Internet and cut and paste loads of articles—reports on Indian Maoists, scientific speculation on whether animals can be gay, critiques of standardized testing—into what typically came to be an eighty-page, single-spaced Word document. Then I would print it out and read it the next day while I was working at the scoring center. This was the only way to avoid going insane. I still managed to score at the average rate for the room and perform according to “quality” standards. While scoring from home, I routinely carry on three or four intense conversations on Gchat. This is the reality of test scoring.

There is a common fantasy that test scorers have some control over the grades they are giving. I laugh whenever someone tells me, “Make sure you go easy and give the kids good grades!” We are entirely beholden to and constrained by the standards set by the states and (supposedly) enforced by the test-scoring companies. To ensure that test scorers are administering the “correct” score, we receive several hours of training per test, and are monitored through varying quality control measures, such as random “validity” papers that are pre-scored and that we must score correctly. This all seems logical and necessary to ensure impartiality—these are, after all, “standardized” tests. Unfortunately, after scoring tests for at least five states over the past three years, the only truly standardized elements I have found are a mystifying training process, supervisors who are often more confused than the scorers themselves, and a pervasive inability of these tests to foster creativity and competent writing.

Scorers often emerge from training more confused than when they started. Usually, within a day or two, when the scores we are giving are inevitably too low (as we attempt to follow the standards laid out in training), we are told to start giving higher scores, or, in the enigmatic language of scoring directors, to “learn to see more papers as a 4.” For some mysterious reason, unbeknownst to test scorers, the scores we are giving are supposed to closely match those given in previous years. So if 40 percent of papers received 3s the previous year (on a scale of 1 to 6), then a similar percentage should receive 3s this year. Lest you think this is an isolated experience, Farley cites similar stories from his fourteen-year test-scoring career in his book, reporting instances where project managers announced that scoring would have to be changed because “our numbers don’t match up with what the psychometricians [the stats people] predicted.” Farley reports the disbelief of one employee that the stats people “know what the scores will be without reading the essays.”²²

I also question how these scores can possibly measure whether students or schools are improving. Are we just trying to match the scores from last year, or are we part of an elaborate game of “joking the stats,” as it’s called on HBO’s *The Wire*, when agents alter statistics to please superiors? For these companies, the ultimate goal is to present acceptable numbers to the state education departments as quickly as possible, beating their deadlines (there are, we are told, \$1 million fines if they miss a deadline). Proving their reliability so they will continue to get more contracts.

As Farley writes, “Too often in my career the test results we returned had to be viewed not as exemplars of educational progress, but rather as numbers produced in a mad rush to get things done, statistics best viewed solely through the prism of profit.”³ It seems to me that what the companies would tell us, if they were honest, would be something like, “Hey guys, your scoring doesn’t really matter. We just want to give the same scores as last year, so that there’s no controversy with the state and we get more contracts and make more profits—so no matter what you learned in training, just try to forget it.” States and local governments, meanwhile, play their own version of this game, because it looks good for them when politicians can claim that test scores are going up. Witness the recent controversy in New York City, where the percentage of students passing the math exam rose from 57 percent in 2006 to 82 percent in 2009, before plummeting back down to 54 percent in 2010 (along with a 43 percent passing rate in English) after the standards were reviewed.⁴

As test scorers, we never know what the numbers we are assigning to papers mean, or where we fit in this elaborate game. We are only responsible for assigning one score, on one small part of a test, and we do not even know whether the score we assign is passing or failing—that information is never divulged in training. We never hear how the students fared. Whether Marissa will be prevented from going to seventh grade with her friends because one of us, before our first cup of coffee kicked in, decided that her paper was “a little more like a 3 than a 4,” we will never know. Whether Marissa’s school will be closed or her teachers fired (to be reborn as test scorers next spring?) remain mysteries to the test scorers. And yet these scores can be of life-and-death importance, as seen in the recent suicide of beloved Los Angeles middle school teacher Rigoberto Ruelas, Jr. Upon learning that he ranked as “less effective” on the *LA Times* teacher performance rating scale—based solely on test scores—Ruelas took his own life.⁵

Even if the scoring were a more exact science, this would in no way make up for the atrocious effect on creativity wrought by the mania

for standardized testing. This impact has now been documented. According to one study, creativity among U.S. children has been in decline since 1990, with a particularly severe drop among those currently between kindergarten and sixth grade.⁶

While test scorers and students might be separated by age, geography, race, and culture, we share one bond: standardized testing puts us to sleep. In the face of the crushing monotony of the hundreds of rote responses fostered by these tests, scorers are left to fight their own individual battles to stay awake. In any test-scoring center, by far the most essential job is done by the person whose sole responsibility consists of making coffee for hundreds of workers, many of whom will consume four to six cups a day to survive. In my mind, I see a hideous symmetry between test scorers' desperate attempts to avoid dozing off, and the sleepy, zombie-like faces of the students as they prepare for and take these tests.

Of course, these students only exist in my imagination. Just as test scorers are never allowed to know the effects of our scores on students, we never get a chance to meet them, to see how they have developed as writers, thinkers, or human beings, or to know what life in their communities or families is like. All we see is a paper on a screen. And after reading hundreds of monotonous papers each day, it's not uncommon to start to feel a bitter distaste for the undoubtedly beautiful youth of America and the seeming poverty of their creative thought.

I remember reading, for twenty-three straight days, the responses of thousands of middle-schoolers to the question, "What is a goal of yours in life?" A plurality devoted several paragraphs to explain that their life's goal was to talk less in class, listen to their teacher, and stop fooling around so much. It's asking too much to hope for great literature on a standardized test. But, given that this is the process through which so many students are learning to write and to think, one would hope for more. These rote responses, in themselves, are a testament to the failure of our education system, its failure to actually connect with kids' lives, to help them develop their humanity and their critical thinking skills, to do more than discipline them and prepare them to be obedient workers—or troops.

While we test scorers might be prone to blame these children for the monotony of their thoughts, it's not their fault that their imaginations and inspirations are being sucked out of them. No points are given for creativity on these tests, although some scorers have told me that, until recently, a number of states did factor creativity into their scores. Ironically, scorers are often delighted to see papers that show individuality and speak in their own voice, and often reward them with

higher scores, though, judging by the papers I've read, it appears as if students often explicitly are told not to be creative. Yet even if creativity were considered, it would not likely do much to change the overall character of the writing—and education—engendered by an emphasis on standardized testing. As Einstein put it, “It is a miracle that curiosity survives formal education.”

An entire education policy that thrives on repetition, monotony, and discipline is being enacted, stunting creativity and curiosity under the guise of the false idol of accountability. What is more, this policy has a differential impact, depending on students' race and class. As Jonathan Kozol explains,

In most suburban schools, teachers know their kids are going to pass the required tests anyway—so No Child Left Behind is an irritant in a good school system, but it doesn't distort the curriculum. It doesn't transform the nature of the school day. But in inner-city schools, testing anxiety not only consumes about a third of the year, but it also requires every minute of the school day in many of these inner-city schools to be directed to a specifically stated test-related skill. Very little art is allowed into these classrooms. Little social studies, really none of the humanities.⁷

Seeing the results of this process is demoralizing to test scorers, and you can feel it in the scoring centers. Even though you can move about freely, use the bathroom when you need, and talk to one another, the room I was in this spring was almost always completely silent. On every project, as the weeks go by, the health of many scorers deteriorates, making me curious as to whether the relentless, soul-crushing monotony of the papers has an actual physical impact on those forced to read them.

To be fair, these papers aren't a total wash. There is often wisdom in them, even on standardized tests. The chasm between rich and poor is at times felt in the writing itself, as some students come from unimaginable privilege, while many more endure heartbreaking experiences in foster homes. The papers are also a testament to the persistence of racism, describing teenagers kicked out of stores or denied service or jobs because of the color of their skin. And it would be wrong to think of test scorers as a down-and-out bunch—many of us do this job in order to avoid having to get other ones that would keep us from our creative endeavors, or from traveling or pursuing other life-enriching possibilities. A number of test scorers I've met over the past three years are authors, artists, photographers, or independent scholars, and it's common to see postings for book releases and other events featuring the work of test scorers on bulletin boards in the break room.

In the error-filled Pearson training video, Marjorie Scardino says, “Most of the people who work at Pearson work with a passion and an intensity, because they think know are doing something important.” But I’ve never gotten the sense from my coworkers that they “think know” what they’re doing is helping kids or the education process. If the Obama administration asked test scorers whether the solution to this country’s education system would be more standardized testing, I think most of them would laugh. I’ve never gotten the sense from my coworkers that they feel that what they’re doing is helping kids or the education process. Unfortunately, the joke is on us, as the Obama administration pushes for even more high-stakes standardized testing. I didn’t know whether to laugh or cry back in April, when all workers at my test-scoring center were asked to fill out a form allowing the company we were working for to get a tax break for hiring us. This tax break came via the Obama administration’s HIRE Act, which was supposed to provide subsidies for companies “creating jobs.” Never mind that we were all going to be hired anyway, because this is seasonal employment. Or that this money was subsidizing temporary jobs with no health care and no hope for transitioning into long-term employment—jobs which, in a better world, would not exist.

While these companies brazenly collect what can only be described as corporate welfare checks, hundreds of thousands of teachers are being laid off, as governments cut funding to education. Maybe next year, some of them will get paid \$12 an hour (or \$10, if they flood the market) to score tests taken by students stuffed into even bigger classes, and help “impartially” decide which schools will be shut down, and which of their colleagues will be laid off. Equally bad, the fanaticism surrounding accountability via testing, which claims it will result in higher-quality teachers, is doing nothing of the sort. Referring to the test-intensive No Child Left Behind Act, Kozol says, “By measuring the success of teachers almost exclusively by the test scores of their pupils, it has rewarded the most robotic teachers, and it’s driving out precisely those contagiously exciting teachers who are capable of critical thinking who urban districts have tried so hard to recruit.”⁸

As a friend of mine was saying his goodbyes to the coworkers in his room at the end of this year’s scoring season, his seventy-year-old supervisor, a veteran test-scoring warrior, uttered the words I imagine many test scorers hope to hear: “I hope I never see you here again.” This is a measure of the cynicism with which many test scorers approach the industry, recognizing that it is fundamentally a game, which too many people are forced to play—but “hey, it beats working at McDonald’s or Subway!”

Yet amid all the hopes of escaping the industry, these test-scoring companies are successfully expanding and are now hoping to get their hands on billions in “school turnaround” money handed out by the Obama administration and state governments. Pearson, for example, has “formed the K-12 Solutions Group, and . . . is seeking school-turnaround contracts in at least eight states . . . [claiming it] could draw on its testing, technology and other products to carry out a coherent school-improvement effort.”⁹

The big test-scoring companies will undoubtedly be called on to furnish their supposed “expertise” in developing and scoring the new generation of more complex tests envisioned by Secretary of Education Arne Duncan. The Obama administration just gave two groups of states \$330 million in grants to develop these new national tests, with the stated aim of assessing more critical thinking skills and providing better feedback to students and teachers. But rather than addressing the problems outlined above, it seems more likely that this move will only transfer the absurdities in current state tests to a national level, with the danger that they will take on an even greater legitimacy. In fact, given that Duncan’s proposal involves even more tests, it is likely to make matters worse.

If scoring is any indication, everyone should be worried about the logic of putting more of our education system in the hands of these for-profit companies, which would love to grow even deeper roots for the commodification of students’ minds. Why would people in their right minds want to leave educational assessment in the hands of poorly trained, overworked, low-paid temps, working for companies interested only in cranking out acceptable numbers and improving their bottom line? Though the odds might seem slim, our collective goal, as students, teachers, parents—and even test scorers—should be to liberate education from this farcical numbers game.

Notes

1. Todd Farley, *Making the Grades: My Misadventures in the Standardized Testing Industry* (San Francisco: Polipoint, 2009).
2. Todd Farley, “A Test Scorer’s Lament,” *Rethinking Schools* (Winter 2008/2009).
3. Todd Farley, “Standardized Tests Are Not the Answer: I Know, I Graded Them,” *Christian Science Monitor*, October 28, 2009.
4. Sharon Otterman, “Confusion on Where City Students Stand,” *New York Times*, August 28, 2010.
5. Alexandra Zavis and Tony Barboza, “Teacher’s suicide shocks school,” *Los Angeles Times*, September 28, 2010, <http://articles.latimes.com>.
6. Po Bronson and Ashley Merryman, “The Creativity Crisis,” *Newsweek*, August 10, 2010.
7. Matthew Fishbane, “Teachers: Be subversive (Interview with Jonathan Kozol),” *Salon.com*, August 30, 2007.
8. *Ibid.*
9. Sam Dillon, “Inexperienced Companies Chase School Reform Funds,” *New York Times*, August 9, 2010.