

Confessions of a Cheacher Tom White



What is the point of teacher evaluation? More specifically, when it comes to measuring student growth, should teachers do things in their classroom and with their students that will result in scores going up, and by extension, increases the likelihood that their own evaluation will look better? Or should teachers do those things - and only those things - that have a real impact on student learning?

It was in the context of these questions that I reflected on a major change in the way I used data in my classroom.

It began last year when I had some time on my hands. I decided to teach a mini STEM unit to my fourth graders. I brought in a box of K'nex parts, including motors and wheels. I divided my students into groups of three and challenged them to build the fastest car. We all started with the same basic design and timed each car three times, using the average speed as a baseline. We agreed on a very strict testing protocol: one person in the group would position the car with its back wheel on one edge of a six-foot table. We then measured the time it took for the front of the car to reach the other end. After each trial, students could change only one aspect of the car; then time it to see if there was an improvement. If there was, they could change something else; if there wasn't, they had to revert to the previous design and try again.

At one point, a group of girls added a "spoiler" to the front of their car. Spoilers, as I understand them, are supposed make a car more aerodynamic by reducing the amount of air that flows under the vehicle. While this makes sense if your car is moving 50, 60 or 100 MPH, it doesn't make much sense for a toy car moving at 2 MPH. The girls' car, however, had a noticeably faster time.

Word spread, and soon every car was equipped with front spoilers. We talked about why the spoilers were having such a drastic effect on car speed. The kids were convinced that it had something to do with aerodynamics. One group of boys, however, finally figured it out. And then they secretly built a five-foot long car and brought it up for its time trial. As you can imagine, it didn't take long for the front of a five-foot long car, with its back wheels at one end of a six-foot long table, to travel to the other end of the table.

The rest of the class cried foul, but the group of boys argued successfully that they did nothing wrong. They simply exploited the rules to their advantage, and even though their car wasn't technically "faster," it did turn in a "faster time."

This seemed like a good place for me to end the unit and move on to something else. But not before I concluded that the boys with the five-foot car did, in fact, win the competition.

What choice did I have? After all, this activity was happening right around the time when I was administering their third, and final, reading comprehension assessment. And my strategy for administering those tests was not unlike that group of boys.

Reading comprehension, of course, is king in fourth grade. Fourth graders usually have decoding figured out, but they still need to lock down their ability to read something and

Tom White has taught elementary education for the past thirty years. Currently he teaches fourth grade in Lynnwood, Washington. He also writes for two education blogs: storiesfromschool.org, which deals with education policy and CORElaborate, which deals with implementation of the common core.



extract every piece of information, both explicit and implicit. In my district we test all elementary students in reading comprehension three times a year.

Not only does every fourth grade teacher take reading comprehension very seriously, but we meet several times as a team, along with the principal, the psychologist, and the learning support teachers to look at reading scores throughout the fourth grade. And although it's not a competition per se, it's definitely an occasion when you want to "look good." And looking good means improving; showing test scores that get better over the course of time.

One of the best ways to show improvement is to start with low scores. This is fairly easy to achieve, as long as you know something about the factors that contribute to high test scores. I know, for example, that kids tend to perform better on Tuesdays and Wednesdays. They tend to do better after they review test-taking strategies. They do better when they receive feedback along the way. They do better when they're encouraged to do better and believe that high achievement is important to them. Of course, they also tend to do better after they've receive quality instruction.

If you know how to get kids to perform to their potential, then you also know how to get kids to perform below their potential. And that's basically what I did, at least on the first test. I gave it on the Thursday and Friday of the first week of school. I spent no time at all reviewing test-taking strategies. I gave no feedback along the way; scoring their work only after they finished. And of course I didn't stress to them how important it was that they do their best, since it actually wasn't their best that I wanted.

In January, I took it up a notch. I gave the test earlier in the week, after reviewing some fairly obvious test-taking strategies. I also told them the scores would count on their report cards, which was true.

Of course I also delivered quality instruction between September and January, and also from January to June, when my class took their third test. And when they took that test, I pulled out all the stops: we tested on Tuesday and Wednesday, reviewed test-taking strategies, and I gave them feedback along the way. After each of the seven parts, they came up to my table where I corrected their efforts, followed by encouragement or perhaps a gentle reminder to "slow down and look for the answers in the text."

As you can probably guess, my class showed a great deal of improvement. They started out low, got better in January and even better in June. And while it would be nice to think that this improvement was due to my excellent instruction, I have no doubt that it was partly due to my manipulation of the testing experience.

Was this cheating? Perhaps, but I don't think so. Or at least I'm not ready to admit it. But it definitely wasn't teaching. I would put it somewhere between the two. Let's call it "cheaching."

Cheaching is bad. Not evil, but bad. Cheaching defeats the purpose of teaching without actually breaking the rules. But before you judge me too harshly, keep in mind that what I do isn't much different than a business executive who times the company's billing cycles in such a way that the earning reports look good to the shareholders. And for that matter, it isn't much different than a trial attorney who chooses witnesses based on how innocently they'll portray the client.

“Cheaching is bad.
Not evil, but bad.
Cheaching defeats the purpose
of teaching without actually
breaking the rules.”

And as I'm sure you can guess, I'm not the only cheacher in the school. Or the district. Or the state. Cheaching is a way of life when you're a teacher. We want to look good, so we look good. It certainly isn't all we do, but it's definitely something we do when we want to look good. And we do want to look good.

Just to be clear, I was cheaching when it didn't even matter. There was nothing at stake other than my own pride. This year, as you can probably guess, I chose to use my students' reading comprehension scores as the basis of my student growth evaluation. And sadly, as you can also probably guess, I had every intention of cheaching. And so I conducted my September assessments exactly the way I did them last year.

But then something happened.

You see, last year at my school we tried something different. Instead of balancing out the eight fourth graders who have IEPs, putting three in my room, three in another room, and two in the third room; we put them all into my room. We did this for a very simple and practical reason. Like most schools,

our special education kids usually get pulled out for special services. Colleen, the learning support teacher, has a separate room, and all the kids leave their regular room for a half-hour or so to work with her. The trouble is, kids coming from different rooms are leaving different classrooms; each classroom is working on a different reading skill or at least working on a different lesson related to that skill. And that makes things messy.

What we decided to do was have all eight of those kids leaving the same classroom, a classroom in which they were all working on the same skill and the same lesson. So we put them all into my room. Colleen pops in at the tail end of my mini-lesson, so she can see whatever skill we are working on and listen to the examples I'm using to reinforce that skill. At the end of the mini-lesson, she'll grab some of the materials, lead six or so kiddos back to her room and leave me with the rest, which I then rotate through my reading groups and independent reading.

It works beautifully. We engage in true collaboration - co-teaching, really. She knows exactly where I'm at and I know what she's doing with my lower readers. We have conversation about those kids nearly every day and we work together to pick relevant reading material for them. We also talk about which kids she should be serving. We use data to make those decisions. When we disaggregate test scores to see which skills are causing the most trouble for certain kids, we can adjust who gets pulled out into Colleen's room and who stays with me.

Of course, collaborating in this way requires reliable data, data that hasn't been doctored. So when the mid-year reading comprehension tests came along, I made sure the testing conditions were as pure as possible: no cheaching tactics at all. I prepared them for their test, gave it to them under optimal conditions and let the chips fall where they may.

I came to regret the approach I took to those tests my students took in the fall. The data simply wasn't good enough to use. Not only that, but I began looking at the data in a different way. It became "our data," not "my data." It belongs to Colleen, our students and me. If it looks good, fine, if it looks bad, then let's do something about it.



Take Lynn, for example, who is a fairly strong reader. Or at least strong enough to keep her off our watch list. However, when I drill into the data, I find that she only got one out of four questions right in the area of distinguishing fact from opinion. So guess what? When our class as a whole starts working on that skill, Lynn will go to Colleen's room, along with Mia, Keenan, Andy, Collin and Dora, who all scored low in that skill. They'll work with her for three weeks, getting the focused instruction they need, until we turn to a new comprehension skill, at which point the kids who need help in that area will get it.

Using data this way is fun. Way more fun than cheating. And more than that, it's effective. We're using data to address our students' needs in a way that's responsible and flexible.

There's only one problem. Beginning this year, the stakes went up. This year, the State of Washington implemented the Teacher and Principal Evaluation Project, or TPEP. One of the requirements of this new teacher evaluation system is that teachers show evidence of student growth—in their whole class, in a subset of students and as a result of collaboration with colleagues.

This means that there's a lot of pressure on everyone to make those scores go up. To the casual observer – or the lawmaker – it all makes sense: better teaching yields better scores. It's as simple as that. Ironically, though, I've learned that I'm less effective when I focus on merely making the data look good. And I'm a lot more effective when I focus on collaboration and using student data to make instructional decisions. The problem, however, is that when there's pressure to produce "good" data, the data we end up with isn't useful for instruction decisions. And that's why we should focus on using data instead of producing it.

I'm convinced that's the direction we should be headed. We, as a system, should turn away from trying to generate good data and make a deliberate shift toward using that data to make responsible, informed instructional decisions. Of course, for that to happen, the shift has to be reflected in our teacher evaluation system. As I see it, the change would be relatively simple: in place of the indicators that call for growth in student data, add indicators that measure how well a teacher uses that student data for instruction.

After all, the whole point of teacher evaluation, as I understand it, is better teaching and more learning.

And less cheating.

Watch Tom's video... <http://tpep-wa.org/student-growth-case-studies>

“ We, as a system, should turn away from trying to generate good data and make a deliberate shift toward using that data to make responsible, informed instructional decisions. ”



Center for Strengthening
the Teaching Profession



@WACSTP
and
@storiesfromschool



253-752-2082
PO Box 7052
Tacoma, WA 98417
www.cstp-wa.org