

[Teacher Blogs](#) > [View From the Cheap Seats](#)

ESEA Hearing: What Wasn't Answered

By Peter Greene on January 26, 2015 9:32 PM

The first Senate hearing on the NCLB rewrite focused on testing and accountability. Discussion at and around the hearing has centered on questions of the Big Standardized Test. How many tests should be given? How often should the test be given? Should it be a federal test or a state test? Who should decide where to draw the pass-fail line on the test?

These are all swell questions to ask, but they are absolutely pointless until we answer a more fundamental question:

What do the tests actually tell us?



Folks keep saying things such as "We need to continue testing because we must have accountability." But that statement assumes that tests actually provide accountability. And that is a gargantuan assumption, leading Congress to contemplate building a five-story grand gothic mansion of accountability on top of a foundation of testing sand in a high stakes swamp.

The question did not go completely unaddressed. [Dr. Martin West led off](#) with some observations about the validity of the test. And then he trotted out Chetty, Friedman and Rockoff (2014) a study that piles tautology (we define good teachers as those with good test results, and then we discover that those good teachers get good test results; also, red paint is red) on top of correlation dressed up as causation. If you like your Chetty debunking with a more scholarly flair, [try this](#). If you like it with Phineas and Ferb references, [try this](#).

Then West piled up more correlation dressed as causation. Citing Deming et al (2014), West takes a stand for the predictive power of testing, and in doing so, he himself makes clear why his support of testing validity is actually no support at all.

Predictive power is not causation. Let's take a stroll through a business district and meet some random folks. I'll bet you that the quality of their shoes is predictive of the quality of their cars and their homes. Expensive shoes predict a Lexus parked in front of a five story grand gothic mansion.

It does not follow, however, that if I buy really nice shoes for all the homeless people in that part of town, they will suddenly have expensive homes and fancy cars.

And here's how test-based accountability works. People off in some capital tell local authorities, "We want to end homelessness. So we expect pictures of all your homeless wearing nice shoes. And if the number doesn't go up, we will dock your pay, kill your dog, and take away your dessert for a year." The local authorities will get those pictures (even if they have to use fake shoes or the same shoes on multiple feet), send off the snapshots to the capital, the capital folks will congratulate themselves for ending poverty, and the homeless people will still be sleeping under a bridge and not in a fancy gothic mansion.

Another version of the same central question that was neither asked nor answered at the hearing would be:

What would give us the best, most complete, most accurate sense of how well educated a young person might be? How many people would seriously answer, "Oh, given the need to measure the full range of a person's skills, knowledge and aptitudes, I would absolutely depend on a bubble test covering just two thin slivers out of the whole pizza of that person." When you think of a well-educated person, do you automatically think of a person who does really well on standardized tests of certain math and reading skills?

Oddly enough, it was a nominally pro-test witness whose testimony underlined that. [Paul Leather, of the New Hampshire Department of Education, testified at some length about the granite state's extensive work](#) in developing something more like a whole-child, full-range assessment-- something that is robust and flexible and individual and authentic and basically everything that a standardized mass-produced

test is not.

Congress put the cart not only before the horse, but before the wheels came back from the blacksmith shop. What they need to do is bring in the testing whizzes of Pearson/PARCC/SBA/etc and ask them to show how the Big Standardized Test measures anything other than a student's ability to take the Big Standardized Test. And I have not even addressed the question of whether or not the Big Standardized Test accurately measures even the slim slice of skills that it claims to assess-- but that question needs to be asked as well. We're missing **serious discussions of testing's actual results, like this one**. Instead, Congress engaged in a long discussion of how best to clean and press the emperor's new clothes.

There is no point in discussing what testing program best provides accountability if the tests do not actually measure any of the things we want schools to be accountable for. You can build your big gothic mansion in the swamp, but it will be sad, scary and dangerous for any people who have to live there.

Tags: [ESEA](#) [testing](#)