

THE CHRONICLE

of Higher Education

Teaching

[Home](#) » [News](#) » [Faculty](#) » [Teaching](#)

July 18, 2010

Education Schools Are Scrutinized for Graduates' Success as Teachers

Huge state databases could transform the assessment of teacher training, but some say the systems have bugs

By David Glenn

Gerald B. Carlson has been dean of the College of Education at the University of Louisiana at Lafayette since 2001. The state's public schools are famously troubled, and Mr. Carlson spends many hours worrying about the quality of his programs. When novice teachers walk out the door at Lafayette, are they actually prepared to teach?

Until two years ago, the dean had only indirect evidence—and the news seemed to be good. His graduates almost invariably pass the state's licensure exams, and they have no trouble finding jobs.

"When we survey employers about our graduates," he says, "93, 94 percent of them say that they seem prepared or well prepared."

But in the summer of 2008, Mr. Carlson got a report card from the Louisiana Board of Regents. Using a new longitudinal database, it had analyzed the standardized-test scores of fourth- through ninth-grade students and matched those scores to the institutions that had trained their teachers.

In one subject area, English-language arts, the recent graduates of an alternative-certification program at Lafayette seemed to be less effective than other new teachers in Louisiana. A year later, Mr. Carlson got more bad news about Lafayette's mainstream B.A. program in education.

Other deans of education across the country can expect to hear similar reports soon. Since 2005 the federal government has given states more than \$400-million to build longitudinal education databases. Before long, nearly every state should have the capacity to do the kinds of analysis that Louisiana has done, tracing schoolchildren's academic performance back to their teachers' teachers.

To many observers, these databases are a long-overdue step toward a real understanding of the quality of higher education. "This kind of information can be extremely powerful," says Grover J. Whitehurst, a former official of the U.S. Education Department who

directs the Brookings Institution's Brown Center on Education Policy. "For the first time, you're getting data on whether training is actually affecting teachers in the way that you would expect it would."

But other people say poorly designed analyses might do more harm than good. The standardized tests that underlie these reports are themselves controversial, of course. Beyond that, some education deans worry that states will crunch the numbers in crude ways that misidentify teachers' effects on their students.

"There's much, much more going on that we'd like to know about," says Sandra L. Robinson, dean of education at the University of Central Florida. "What about the progress of the students toward graduation? What about referrals to counseling? There are so many more impacts that a teacher has in the classroom that aren't captured by standardized tests."

Identifying the Culprit

When Lafayette got its report card, Mr. Carlson and his colleagues did not have much trouble finding a plausible culprit: Their alternative-certification program, which was designed to speed students through with just 33 credits, rather than the traditional 128, did not include a course in teaching the language arts. They added such a course last year.

"I was fully prepared to close that program if we couldn't find out what the deficiency was," says E. Joseph Savoie, president of the Lafayette campus. "But we think we were able to come up with the right diagnosis."

There is a small irony in the fact that Mr. Savoie, of all people, had to go through that process. From 1996 to 2007, he served as Louisiana's higher-education commissioner. In that role he pushed hard for new methods of assessing teacher education. Louisiana's new methods of analysis are his child as much as anyone's. But after he became president in Lafayette, his hometown, in 2008, he was one of the first campus chiefs to receive a less-than-average report under the new system.

Lafayette's report card was not actually that bad. Language arts is only one of five subject areas covered in the state's analysis, and even in language arts, measures of Lafayette graduates were only slightly below average. But still, the report stung.

"I was bound and determined to find out what had caused the problem," Mr. Savoie says. "We had meetings with faculty and provosts, and I called George Noell to ask for some additional

analysis."

George H. Noell is a professor of psychology at Louisiana State University at Baton Rouge, and lately his phone has been ringing with such requests.

Seven years ago, Mr. Savoie and other state officials asked Mr. Noell to develop a statistical framework for using students' test scores to assess teacher-preparation programs. He and a few LSU colleagues traveled the state visiting schools of education on sun-baked crossroads.

"Everybody said the same thing to us," Mr. Noell says. "We're fine. Why are you bothering us? We have a great product. We feel like we produce great teachers, and when we survey principals, they say that our teachers are strong.'

"And that's true. The survey data is uniformly positive. And yet when you look at the level of achievement of schoolchildren in Louisiana, things are not good. So there have to be issues somewhere. We came to the realization that without some sort of objective metric, we probably couldn't make any progress, because people were convinced that they were fine."

Despite his occasionally sharp tone, Mr. Noell says he has never wanted to bludgeon schools of education with bad news. Precisely because he wants to persuade deans that their programs can improve, he has worked with them to make sure that the state's reports are accurate and fair. He does not want teacher-training programs to be blamed for factors outside their control.

After a few years of trial and error, Mr. Noell and his colleagues have built a system based on "value added" measures, not raw test scores. Teachers are assessed according to how well they build their students' skills from one year to the next, so they are not penalized for having students who arrived in their classrooms with weak skills. That means, in turn, that teacher-training programs are not punished as a result of a large proportion of their graduates' choosing to work in high-poverty districts.

In addition, the analyses cover a three-year range, so they do not suffer from the statistical volatility that often afflicts single-year measures of teachers' effects. And the Louisiana reports include only students who stay in the same classroom throughout the academic year, so teachers are not unfairly penalized (or rewarded) for midyear transfers.

"We look at this as a very positive process," says Mr. Carlson, the

education dean at Lafayette. "George and his team are always careful with the data, and we've been able to use these reports to improve our programs."

Other Data Questioned

A few hundred miles to the east, education deans are less sanguine about the new era of longitudinal analysis. Last November, Florida's State Board of Education released report cards that ranked the effectiveness of recent graduates of each of the state's 54 teacher-training programs, based on students' test scores in the fourth through 10th grades. The reports were less statistically sophisticated than Louisiana's, and several deans objected that the findings were severely misleading.

"We were all astounded and frustrated," says Ms. Robinson, of Central Florida. "There seemed to be no careful culling of the data."

Her anxiety was especially sharp because Florida's Legislature was debating Senate Bill 6, a controversial measure that was vetoed by Gov. Charlie Crist in April. The bill would have effectively ended tenure for schoolteachers in Florida. A lesser-known provision would have required the state to consider student test scores in reapproving teacher-training programs.

Ms. Robinson says she has no objection to such a requirement, in theory. But the analyses that Florida released last fall were so badly designed, she says, that her heart sank when she thought they might someday have the force of law. She rallied other deans in the state to contest the report. Their complaints included the following:

- Mismatched data. The Florida report analyzed the performance of novice teachers who graduated from teacher-training programs in 2008, based on their students' test scores during the 2008-9 academic year. But a coding error meant that some people who earned nonteaching degrees in 2008 were included in the analysis—for example, someone who received a teaching degree at Florida State University in 1993 and a master's degree in educational leadership (for aspiring principals) from the University of South Florida in 2008. "People like that were included in our report," says Colleen S. Kennedy, dean of education at South Florida. "But we shouldn't be held responsible for their teaching performance if we didn't train them as teachers."
- Apples-to-oranges comparisons. The report compared teacher-training programs as a whole, but some small alternative-certification programs specialize in particular grade levels. Ms. Robinson says it is unfair to compare her program in its entirety with programs that train only elementary-level teachers, because students' test-score gains are often stronger in fourth grade than in 10th grade. (The Louisiana reports suffer the same limitation.)
- Weak statistical power. The Florida report analyzes only a year of data, while the Louisiana reports cover three years. And the Florida report calculated scores for teacher-training programs based on as few as 10 teachers, while the Louisiana system requires at least 25 teachers to be identified before it rates a training program. Because it is based on relatively few observations, the Florida report's reliability is suspect, the critics say.
- Out-of-field assignments. The Louisiana model includes only teachers who are teaching the subjects in which they were certified, but the Florida report has no such control. Larry G. Daniel, dean of education at the University of North Florida, says a significant number of his program's graduates are assigned to teach subjects other than the ones they studied.

"So, in other words, we might be held responsible for someone we had not even prepared to be a math teacher, someone we had prepared to do something else," Mr. Daniel says. "The state really needs to continue to work those bugs out of the system."

Kathryn S. Hebda, the Florida official who oversaw last year's report, says that it was a preliminary model, and that she and her colleagues are indeed refining it. Ms. Hebda, deputy chancellor for teacher quality at the State Board of Education, says Florida will soon adopt a value-added system for measuring individual teachers' performance.

Once that system is in place, she says, evaluations of teacher-training programs will use a more sophisticated statistical model, perhaps more like Louisiana's.

"We have a lot of work to do, but it's exciting work," Ms. Hebda says. Despite the conflict over last fall's report, she adds, "I think all of the teacher-training institutions in Florida agree that we need better ways of measuring performance. They want to be better informed, we at the board want to be better informed, and the public wants to be better informed."

Meanwhile, two scholars at Harvard University released a working paper in June that used student test-score data to compare the performance of 11 large teacher-training programs at Florida colleges and universities. The authors found no strong evidence that any program is significantly better or worse than the others. They also found that teachers who hold master's degrees are no more effective, in general, than teachers who hold only bachelor's degrees. The paper argues that school districts should reward teachers for their actual classroom performance, not for the credentials they carry.

Despite their anger about last fall's report, Ms. Robinson and her fellow deans have not outright opposed longitudinal analyses of student achievement in Florida. Instead they called in Mr. Noell, the LSU psychologist, whom Ms. Robinson met last year in Washington at a meeting with U.S. Secretary of Education Arne Duncan. In February, Mr. Noell met with two dozen leaders of teacher-training programs in Florida. (He says he has also consulted in eight other states recently.)

"I think Noell and his group in Louisiana are quite a bit ahead of us here," says Mr. Daniel, of North Florida. "They've struggled with those issues and dealt with them quite effectively."

Some observers are skeptical even of relatively sophisticated models like Louisiana's. "Any accountability system worth its salt will not only tell you who's doing better or not, but why," says Barnett Berry, president and chief executive of the Center for Teaching Quality, a nonprofit organization in North Carolina.

He doubts that measures based only on high-stakes standardized tests can capture the full complexity of teachers' work. He also worries that the analyses will fail to take into account school-level variables like the mentoring (or lack thereof) that new teachers receive, or the camaraderie and stability of a school's faculty.

But Mr. Whitehurst, of the Brookings Institution, says it would be a mistake to try to feed too many variables into analyses like Louisiana's.

"Academics are expecting more out of a value-added system for evaluating teachers than they would expect or find in any other industry," he says. "The question is, I think, not whether these value-added systems are perfectly reliable, but whether, on balance, they improve performance by doing a better job than the current evaluation systems, which don't exist at all, really. Do they do a better job at identifying persistently good and bad performers?"

Mr. Daniel has no doubt that his programs, at North Florida, will eventually be judged according to some kind of longitudinal data analysis. He just wants the system to be a good one. "Being able to look at what impact a teacher is having on student learning will become more and more important, not less so," he says. "We're at a crossroads right now in our profession."

Copyright 2010. All rights reserved.

The Chronicle of Higher Education 1255 Twenty-Third St, N.W. Washington, D.C. 20037