

January 21, 2008

## New York Measuring Teachers by Test Scores

By JENNIFER MEDINA

New York City has embarked on an ambitious experiment, yet to be announced, in which some 2,500 teachers are being measured on how much their students improve on annual standardized tests.

The move is so contentious that principals in some of the 140 schools participating have not told their teachers that they are being scrutinized based on student performance and improvement.

While officials say it is too early to determine how they will use the data, which is already being collected, they say it could eventually be used to help make decisions on teacher tenure or as a significant element in performance evaluations and bonuses. And they hold out the possibility that the ratings for individual teachers could be made public.

“If the only thing we do is make this data available to every person in the city — every teacher, every parent, every principal, and say do with it what you will — that will have been a powerful step forward,” said Chris Cerf, the deputy schools chancellor who is overseeing the project. “If you know as a parent what’s the deal, I think that whole aspect will change behavior.”

The effort comes as educators nationwide are struggling to figure out how to find, train and measure good teachers. Many education experts say that until teacher quality improves in urban schools, student performance is likely to stagnate and the achievement gap between white and minority students will never be closed. Other school systems, including those in Dallas and Houston as well as in the whole state of Tennessee, are also using student performance and improvement as factors in evaluating teachers.

The United Federation of Teachers, the city’s teachers’ union, has known about the experiment for months, but has not been told which schools are involved, because the Education Department has promised those principals confidentiality.

Randi Weingarten, the union president, said she had grave reservations about the project, and would fight if the city tried to use the information for tenure or formal evaluations or even publicized it. She and the city disagree over whether such moves would be allowed under the contract.

There is no way that any of this current data could actually, fairly, honestly or with any integrity be used to isolate the contributions of an individual teacher,” Ms. Weingarten said. “If one permitted this, it would be one of the worst decisions of my professional life.”

New York invited principals from hundreds of elementary and middle schools with sufficient annual testing data to participate in the program, which will produce an elaborate stream of data on 2,500 teachers.

140 schools — a tenth of the roughly 1,400 in the system — teachers are being measured on how many students in their classes meet basic progress goals, how much student performance grows each year, and how that improvement compares with the performance of similar students with other teachers.

In another 140 schools, principals are being asked to make subjective evaluations of roughly the same number of teachers so officials can see if the two systems produce widely disparate results. New York City schools employ roughly 77,000 teachers. In all 280 schools, the principals agreed to participate in the program.

Deputy Chancellor Cerf said that how students performed on tests would not be the only factor considered in any system to rate teachers. All decisions will include personal circumstances and experiences, he said, but the point will be to put a focus on whether or not students are improving.

“This isn’t about how hard we try,” Mr. Cerf said. “This is about however you got here, are your students learning?”

Ms. Weingarten said the system was not needed. “Any real educator can know within five minutes of walking into a classroom if a teacher is effective,” she said. “These tests were never intended and have never been validated for the use of evaluating teachers.”

The experiment is in line with the city’s increasing use of standardized test scores to measure whether students are improving, and to judge school quality. A new bonus program for teachers and principals, as well as the letter grading system for schools unveiled last fall, are all linked to improvement in scores. Nationally, too, school systems are increasingly relying on these measures to judge schools.

Virtually all education experts agree that finding high-quality teachers is critical to improving student learning, particularly in high-poverty urban areas, where good teachers are usually more difficult to find. Recent research has found that the best teachers can help struggling students catch up to more advanced students within three years.

But experts are grappling with how to determine what makes a good teacher. Neither graduate programs in education schools nor previous academic records are reliable predictors, they say. The federal No Child Left Behind law requires that districts place a “highly qualified” teacher in every classroom, which typically means one who has completed a certification program, but this, too, is not necessarily a good indicator of quality.

It seems hard to know who is going to be effective in the classroom until they are actually in the classroom,” said Thomas J. Kane, a professor of education and economics at Harvard, who is conducting

several research projects on teacher quality in New York City, and who is involved in the new effort.

Mr. Kane said there was little evidence that teachers with the “right paper qualifications” were any more effective than those without them. “But most school districts spend very little time trying to assess how good teachers are in their first couple of years, when it is most important,” he said.

Nationwide, more than 95 percent of teachers receive tenure within their first three years of teaching, according to some studies. And once teachers receive tenure, it is extremely difficult to have them removed from classrooms.

In some sense, New York’s effort to judge teachers partly on their students’ improvement is a logical extension of the grading system for schools that was unveiled last fall, although officials adamantly say they have no plans to assign letter grades to individual teachers.

“I don’t think anyone here would embrace the formulaic use of even the most sophisticated instrument — you get tenure if this, you don’t get tenure if that,” Mr. Cerf said.

He added that the new effort was just one of several ways in which the city was exploring how to evaluate and improve teacher quality. In recent months, city officials have begun training new lawyers to help principals navigate the considerable red tape required to remove inadequate teachers.

They have increased recruiting efforts to attract talented teachers to hard-to-staff schools. And they are allowing schools to earn merit bonus pools to distribute to teachers based on test scores.

“This should simply be one more way to think about things,” said Frank A. Cimino, the principal of P.S. 193 in Brooklyn, who said he was participating in the experiment. “It is going to tell you some things you don’t know, but it will miss the other things that go on in a classroom.”

William Sanders, a researcher in North Carolina who was one of the first to begin evaluating teachers and schools based on student test score improvements, said that while such a system could be used to make broad judgments, it was difficult to use it with precision enough to find differences among teachers who are simply average.

“Can you distinguish the top teachers? Yes,” Mr. Sanders said. “Can you distinguish the bottom teachers? The answer is yes, too. But it would be risky to make decisions using information at the classroom level for teachers who are just in the middle. You might miss a lot that way.”

The city’s pilot program uses a statistical analysis to measure students’ previous-year test scores, their numbers of absences and whether they receive special education services or free lunch, as well as class size, among other factors.

Based on all those factors, that analysis then sets a “predicted gain” for a teacher’s class, which is measured

against students' actual gains to determine how much a teacher has contributed to students' growth.

The two-page report for each teacher examines information both from one year and over three years. The information also compares the teacher with all other teachers in the city, and with teachers who have similar classrooms and experience levels. The second part of the report measures how well a teacher does with students with different skill levels, showing, for example, whether the teacher seems to work well with struggling students.

Mr. Cerf said officials expected to decide by the "early summer" whether they would use the analysis to evaluate individual teachers for tenure or other decisions, and if so, how they would do so. Such a decision would undoubtedly open up a legal battle with the teacher's union.

Copyright 2008 The New York Times Company

[Privacy Policy](#) | [Search](#) | [Corrections](#) | [RSS](#) | [First Look](#) | [Help](#) | [Contact Us](#) | [Work for Us](#) | [Site Map](#)

---