

CEPA

CENTER FOR EDUCATIONAL POLICY ANALYSIS

In This Issue

This newsletter summarizes findings from three years of research on the influences of New Jersey's Elementary School Performance Assessment (ESPA) on the teaching of fourth grade mathematics and science. Based on a statewide survey, in-class observations, and teacher interviews, we conclude that:

- The changes in content taught have been very modest, most notably a reduction in the amount of teaching of basic mathematical operations.
 - The presence of performance or open-ended items on the test has sensitized teachers to new, more challenging instructional strategies, but that practice has not yet changed appreciably.
 - Teaching to the test includes a mix of introducing more intellectually challenging practices—a positive outcome—and narrowing instruction, a negative one. The amount of teaching to the test has not changed over three years, but there is more use of commercial test preparation materials in the poorest districts which suggests that the tests are having negative consequences for the state's poorest students.
 - Teaching changes more where principals are supportive. Because principals' support is so general and rarely reflects a strong view on instructional content or methods, however, some of the resulting changes reflect state standards while others may work against those standards.
- We recommend the following changes for the future:
- As the state moves to a new testing system, it is critical to maintain performance assessment items in the test.
 - Changing tests won't improve teaching without more direct attention to deep and sustained professional development for teachers.
 - Professional development to help principals understand state standards and new approaches to instruction is also very important.

Much Ado About Testing: The Effects of Three Years of State Testing on Teaching

New Jersey began administering its fourth grade assessment—the Elementary School Performance Assessment (ESPA)—in the spring of 1999. Now with a new governor in office and new federal legislation requiring more testing, the state is rethinking its whole assessment system.

To ensure that we learn from the last three years' experience, we report on a longitudinal study of how state testing has affected teaching. Our conclusion, very simply stated, is that there were fewer changes than might have been expected. After three years of experience, 4th grade math and science courses cover essentially the same topics in both math and science using very close to the same instructional approaches as before. Teaching to the test is present, but not growing. One thing that did help to improve teaching was the performance assessment items on the test. Another was support from building principals, but the differences were not as striking as might have been hoped. As the state moves forward, we have three recommendations. The state should:

- 1 Keep the performance items now used in state tests in any future test.
- 2 Move beyond just changing tests to helping teachers improve their teaching approaches in deep and sustained ways.
- 3 Increase focus on building principals' knowledge about effective, standards-based teaching and how to help teachers develop it.

Content Coverage

ESPA is intended to be aligned with New Jersey's Core Curriculum Content Standards. In math and science, these standards are consistent with national standards which call for modernizing the content taught. National standards also suggest using instructional methods that feature student inquiry with the expectation that inquiry will help students develop a deeper understanding of content with no loss of facility in basic mathematical operations. For most educators, the most concrete aspect of the standards is the state testing program. The change in standards and tests raises two questions. First,

continued on next page

Many teachers reported that ESPA is encouraging them to implement more inquiry-oriented instructional practice, and most view such changes positively.

continued from page 1

are teachers teaching different content? Second, are they using more standards-based instructional approaches?

Given the complaints in other states about how teaching to the test changes curriculum coverage, content changes were remarkably modest. In the spring of 1999, we began surveying 4th grade teachers throughout the state (see box on study methods). This survey was repeated in the spring of 2000 and 2001. Each year, we asked teachers how much time they spent teaching 17 different mathematics topics. The teachers who participated in all three years of the survey did not report any change in the time they spent on thirteen topics. They reported changes in four areas. The amount of time declined in two areas related to teaching such whole number operations as addition, subtraction, multiplication, and division (See Figure 1). Interviews confirmed that teachers did reduce time on basic operations. This reduction would have been helpful if teachers reported a comparable increase in some of the areas suggested by the state and national standards (such as data analysis and algebraic reasoning). Unfortunately, we did not see such changes. The only increase we did find occurred in the use of open sentences and variables. Time spent on statistics increased between 1999 and 2000 but then declined.

In science, we saw a three year decline in five topics with most of the decline between 1999 and 2000 (See Figure 2). This is somewhat misleading. Our interviews suggest that teachers increased the amount of science they

taught in preparation for the new tests that started in 1999. However, we suspect that because the percent of students passing the science tests was quite high statewide, higher than on the language arts portion, attention shifted to other areas.

In sum, the 4th grade test reduced the pressure to teach basic mathematical operations but did not encourage the introduction of new math topics or sustain an interest in science.

Instruction

In interviews, teachers report that they are rethinking their instructional strategies in math and science, but our observations indicate that they are not making substantial changes. Many teachers reported that ESPA is encouraging them to implement more inquiry-oriented instructional practice, and most view such changes positively. They emphasize that these changes result from the inclusion of performance assessment or open-ended items on the test. These items require students to show their work and justify their response. In this regard, the teachers note that performance items provide an impetus to include some of the more challenging instructional practices New Jersey's new standards are supposed to promote. One teacher explained that as a result of these new items, "It's become my philosophy to teach them the concepts before, just, you know, ramming these rote facts down their throats."

Teachers mentioned several changes they were trying to incorporate. Two-fifths talked about trying to get students to explain their thinking in more detail. According to one teacher, the part "that I

guess I really didn't do a lot of before is really get the students to start to learn how to explain their thinking, to explain what they were doing. Sometimes they do it in writing; sometimes they do it to a partner; sometimes they do it to me." One strategy to encourage student explanation is the use of more open-ended questions on tests and in class, and was mentioned by a third of the teachers.

Almost half the teachers said they used manipulatives or concrete objects to help students explore mathematical ideas. The ESPA has questions that involve written or pictorial descriptions of manipulatives, and many teachers felt that students who are more familiar with some of the current manipulatives would respond better to those questions.

Forty percent of the teachers said they had students spend more time on writing in mathematics. One teacher said that she now had her students "write all the time for all subjects." Some teachers used "writing" as one way to have their students explain their line of thinking in mathematics. In fact, a fifth of the teachers said they had students keep journals in math, as well as other subjects.

When we went into classrooms, however, changes appeared smaller than the interviews suggested. For instance, as the interviews indicated, manipulatives were used in two-fifths of all observed lessons. The problem is that they were often used to reinforce older, drill-based instructional approaches, rather than the newer more inquiry-oriented strategies suggested by national standards. In fact, in almost two thirds of the lessons where manipulatives were used, the teacher told the students

Figure 1: Changes in Math Content

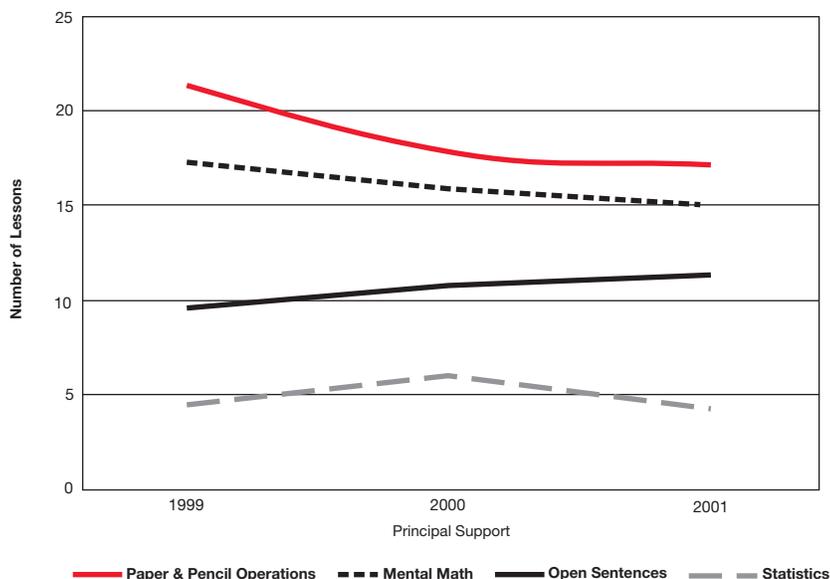
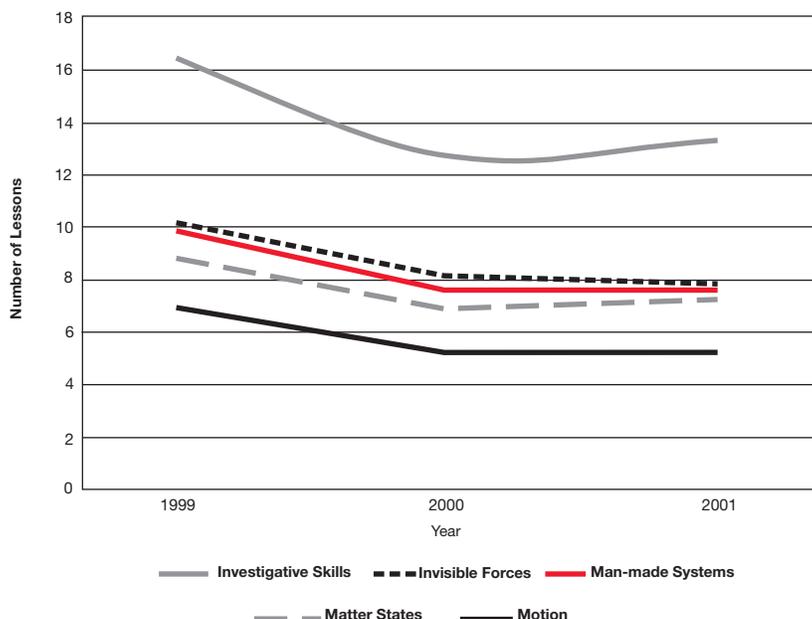


Figure 2: Changes in Science Content



What is striking is how little change there has been in test preparation.

exactly what to do with the materials, and the students did it as best they could. Teachers let children use manipulatives to explore mathematical ideas in deeper ways in only a fifth of the observed lessons.

Beyond looking at the use of specific practices and materials, we also examined the mathematical tasks students were asked to perform. We categorized tasks as involving memorization only; doing procedures where the focus was on getting the right answer without understanding; doing procedures to develop a deeper understanding of mathematical ideas; or, doing a mathematical task that requires complex thinking consistent with the “spirit” of the standards. Three-fourths of all tasks fell into the first two categories (memorization or doing procedures in an algorithmic manner). Only one out of forty observed lessons gave students the chance to think in more complex ways.

Most teachers we observed are adopting new strategies, like using manipulatives and small group instruction, as part of their instructional practice. However, classroom discussion rarely encourages students to communicate their thinking or justify their solutions. The teachers we talked to said that they have been motivated to change their teaching styles as a result of the performance assessment or open-ended items on the ESPA, but our observations suggest that these changes remain superficial, and not of the quality implied by state and national standards.

Limited change is to be expected from changing a test. New test items provide examples of what different kinds of instructional tasks might be, but they

don't help teachers understand how to use those tasks, what kinds of questions students might ask, or the mathematical content of such tasks. For that, they need more help than a test can provide.

Teaching to the Test

One fear with state testing is that it will lead to teaching to the test, where teachers prepare students to score high, but not necessarily to understand the underlying content. Teaching to the test is generally seen as problematic for several reasons: topics that are not tested may not be taught; when topics are taught, the emphasis is on learning how to get the right answers, not learning to understand the ideas and methods involved; and teaching to the test may “pollute” the test (that is, test scores may reflect the special tutoring, not children's underlying knowledge of the field). On the other hand, teaching to the test may be a part of a reform process. The growing use of manipulatives and writing in mathematics already described are examples of practices that are a part (but only a part) of an overall strategy consistent with recommendations of national experts.

We developed a scale for teaching to the test that included a variety of practices, including teaching test-taking mechanics (i.e. how to put your name on the test); teaching test-busting skills like methods for turning story problems into arithmetic calculations; and using commercial test-preparation materials. Most of these activities narrow the curriculum, reduce math and science to a limited number of calculation tricks, or devote time to the test itself rather than

the subjects to be tested. Teachers were asked how much test preparation they did the month before the test and throughout the year.

What is striking is how little change there has been in test preparation practices. Data from the three year survey show that teachers did about the same the amount of teaching to the test in 2001 as in 1999.

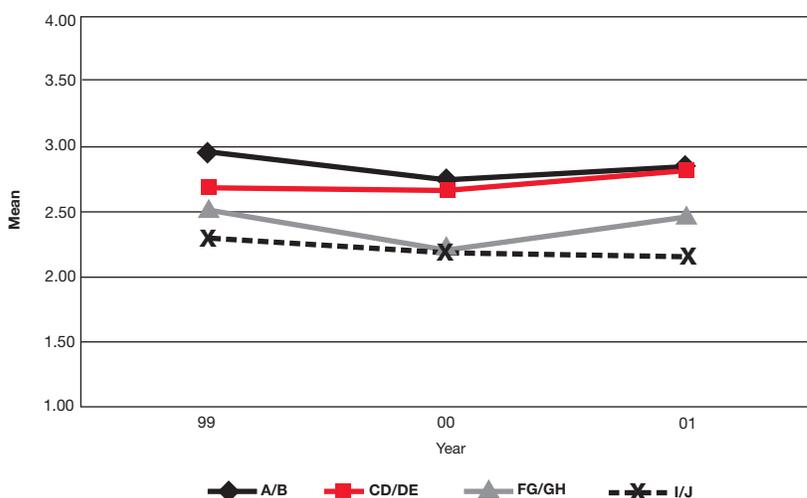
We also found very few differences when we examined specific test preparation practices. The most important did not concern differences over time but between the richest and poorest districts. The state of New Jersey classifies school districts into “District Factor Groups” running from A (the poorest) to J (the wealthiest). The poorest districts consistently use commercial test preparation materials more than do wealthier districts. Figure 3 shows this difference for the month before the ESPA for four groups of DFGs. Every year, the poorest districts used commercial materials the most and the richest districts used them the least, with the two other groups falling in between. The greater use of test preparation materials in poor districts probably reflects the strong pressures in these districts to increase their test scores.

Last year, we reported that at least one more positive test preparation practice was high in both the richest and poorest districts and lowest in the middle. That finding was not replicated this year. On balance, then, test preparation practices seem to be undermining more challenging approaches to math and science the most in the poorest districts.

Leadership

Although we saw very little change in instructional strategies, there was more change where principals supported teachers. However, this change was not particularly focused. The teacher survey includes three measures of teaching practice. In addition to the teaching to the test measure, we asked teachers about their use of inquiry-oriented instruction. This approach reflects the aspirations of standards-based reformers who call for teaching

Figure 3: DFG Differences in Use of Commercial Test Prep Materials

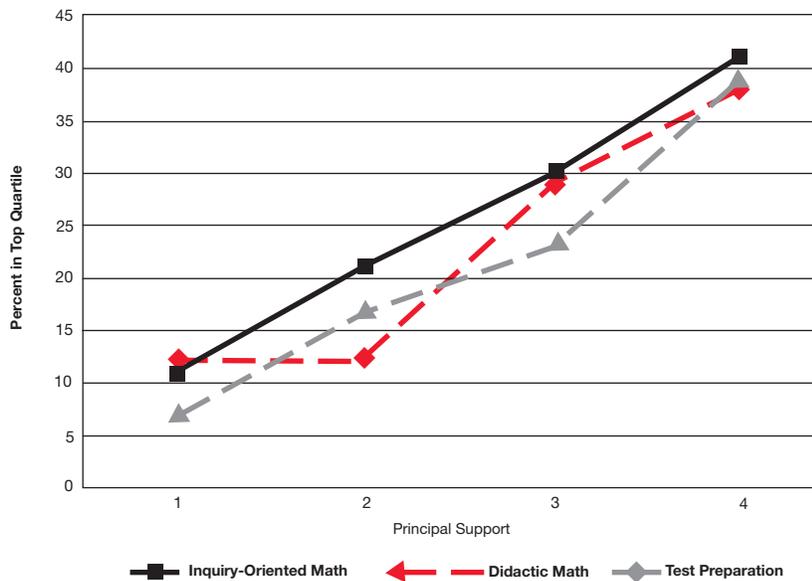


that helps students develop a deeper understanding of these subjects. Students can do so by actively engaging in thoughtful problem solving investigations involving mathematical and scientific ideas and by explaining and justifying their thinking to peers and teachers. This approach is in keeping with New Jersey’s mathematics and science standards. Didactic instruction reflects conventional approaches in mathematics where teachers describe computational

procedures and have students practice them in a structured, repetitive way. In science, teachers explain first and may have students follow highly prescribed experimental procedures. While students may learn facts and procedures from didactic instruction, they rarely develop deeper understanding.

Principal support for standards and standards-based assessments influenced all three dimensions of teaching practice. In fact, teachers who experienced high principal support engaged in more of everything (see Figure 4). This

Figure 4: Principal Support and Training



is somewhat surprising. Many people view inquiry-oriented and didactic instruction as contradictory. Much of the standards reform movement is about helping teachers shift from didactic to inquiry-oriented teaching. And teaching to the test is suspect because it may mean that test scores are going up without students actually learning more math and science.

This situation may reflect both principal pragmatism and limited understanding of the true intent of the standards and curriculum content. Most elementary principals are generalists who want to help their schools perform better and their teachers improve. They rarely have strong opinions about state standards or deep knowledge about how to teach math and science. Thus, supportive principals will be open to many approaches to improving student performance and may look to others for subject-specific knowledge.

The assistance principals provide is very general. In response to testing, almost all principals bought more resources like books and materials for teachers. According to interviewed teachers, what separated the principals who supported standards from the other, was their helping teachers to teach better. Such leadership was usually indirect. Twenty-two of 56 principals described in teacher interviews helped arrange for teachers to get extra training to teach mathematics and science. But 12 of the 15 highly supportive principals did so as opposed to only three of nine who provided very little support. Such support does not include principal-provided training. It does cover situations where principals tell teachers about training opportunities and provide substitutes so they can attend. One teacher described such a principal as “very helpful. He’s allowed us to go to some workshops and he’s given us the freedom to choose some effective professional development opportunities and he’s really been supportive of that in terms of time.”

Supportive principals rarely use formal teacher observation and evaluation. Only 16 percent of principals said that they used observations to ensure that students are learning the core curriculum content standards, and there was a slight tendency for this activity to be mentioned more by the less supportive principals (21 percent) than the most supportive (13 percent). Overall, supportive principals used information, but the information they had was general and rarely reflected a solid knowledge of new instructional approaches. They also used informal influence rather than formal authority to get teachers to change their practice.

New Jersey principals rarely use the exemplary leadership practices that research suggests are most effective. While they work very hard, many are not charismatic individuals who lead through idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration in the service of improvement. Moreover, they do not try to shape the whole school culture to build a shared vision of better teaching for all children. In sum, principals have great potential to improve the quality of teaching, but like teachers, they need learn more about new approaches to teaching to high standards and helping teachers improve.

A Fork in the Road

New Jersey is at a fork in its road to higher standards. The state government is very busy revising its standards and tests. However, our research suggests that we can work on these tasks forever without doing much more to improve

teaching and, ultimately, student achievement. In fact, revising tests can make things worse if the performance items that have been the only part of the state tests to motivate change in practice are eliminated.

To really improve teaching in New Jersey, the state might consider raising the stakes by increasing the rewards and punishments linked to state tests. However, the clues in our data and the experience of other states suggests that doing so is a bad idea. The worst aspects of teaching to the test often increase when the stakes go up. There are enough instances around the country of children being inappropriately held back or sent to summer school because of mistakes in test scoring to question raising stakes. Most national experts agree that test scores should only be used with other supporting evidence when making such high stakes decisions.

Three other strategies are more valuable. First, keep performance items in state tests. They, more than anything else in the current system, raise the issue of teaching to more challenging intellectual standards. Second, move beyond existing policies to provide a sustained, comprehensive effort to help teachers develop their teaching skills. Professional development should be continuous, deep in content and pedagogical knowledge, and reinforced with follow-ups and in-class support. Third, remember the principals. They have great potential to help teachers improve but need to know more about the new standards and how to support educational reform to take advantage of that potential.

About the Study

This newsletter presents findings from a three-year study of the impact of New Jersey's Elementary School Performance Assessment (ESPA). Information comes from two sources. The first is a longitudinal survey of 4th grade teachers. We received responses from 247 teachers in 1999, 287 teachers in 2000, and 301 in 2001. One hundred and nineteen teachers participated in all three years. The samples are representative with respect to district wealth and region. The survey was conducted by the Center for Public Interest Polling at Rutgers University's Eagleton Institute. Survey interviews lasted about 45 minutes.

The second data source was observations and interviews with 10 teachers in 1999, 53 in 2000, and 29 in 2001. Teachers were generally observed in two mathematics and one science class. They were also asked a series of open-ended questions about the lessons observed, their views on ESPA, and opportunities to learn more about the teaching of mathematics and science. Field work was done by the staff of the Center for Educational Policy Analysis at Rutgers University.

The research team for this newsletter included William Firestone, Roberta Schorr, and Lora Monfils. Monfils is a research associate at the Center for Educational Policy Analysis. Schorr is assistant professor of mathematics education at Rutgers University-Newark. Firestone is professor of Educational Policy and Director of the Center for Educational Policy Analysis at Rutgers. We wish to thank Gregory Camilli, Katrina Bulkley, Jennifer Hicks, Cecilia Martinez, Yasuko Munteneau, and Terrie Polovsky and Terri Hawkes for their assistance. This research was funded by grants # 9804925 and 9980458 from the National Science Foundation.

CEPA

Much Ado About Testing: The Effects of Three Years of State Testing on Teaching

MAY 2002

The Center for Educational Policy Analysis is located at Rutgers University's Graduate School of Education. Its task is to analyze how policies are made and implemented, and their effects on children and teachers in order to support more effective policy-making in the future.



THE STATE UNIVERSITY OF NEW JERSEY
RUTGERS

Center for Educational Policy Analysis in New Jersey
Graduate School of Education
10 Seminary Place
New Brunswick, NJ 08901-1183