

. THE IMPACT OF PREPARING FOR THE TEST ON CLASSROOM PRACTICE

Roberta Y. Schorr

Rutgers University

Sylvia Bulgar

Rider University

For the past three years, we have been investigating the implications of testing on students in elementary mathematics classrooms throughout the state of NJ. Our research, and the work of others, indicates that the impact of testing depends on many factors, not the least of which involves how teachers prepare their students for the test. In this particular report we focus on the ways in which teachers engage in what is commonly referred to as “teaching to the test”. Our results indicate that this can manifest itself in many ways ranging from drill on test taking mechanics to the incorporation of more complex problem activities into everyday instruction. We explore how each plays out in the context of classroom instruction, and the implications this may have for students.

BACKGROUND AND THEORETICAL FRAMEWORK

This paper¹ builds upon previous publications in which we discussed the broader implications that testing has had on teaching practices at the fourth grade level throughout the state of NJ² (Schorr & Firestone, 2001; Schorr, Bulgar, Monfils & Firestone, 2002; Monfils et al 2002; Firestone, et. al. 2002; Schorr et al, in press)³. In our previous papers we summarized our findings with respect to the overall changes that teachers were making in their regular classroom practice in response to the test. Briefly stated, we noted that teachers are now increasing their use of hands-on manipulatives, small group instruction and real life problem activities. However, we also reported that these increases have not been accompanied by some of the deeper changes advocated in state and national standards in mathematics (NCTM, 2000; NJ Mathematics Curriculum Frameworks, 1996) including closer attention to children’s mathematical thinking, increased classroom discourse, and an approach to mathematics instruction that is more focused on learning with understanding. We have also reported that many of the teachers believe that they have either introduced or increased their coverage of topics that had been addressed with low historical frequency in the 4th grade math curriculum (such as, probability, statistics and data analysis).

This particular paper explores a related, but different set of issues, namely the specific techniques that teachers use to prepare their students to do well on the state test, and how

¹ The work on this paper was supported by two grants from the National Science Foundation. The opinions presented here are those of the authors and are not necessarily shared by NSF, Rutgers University or Rider University.

² Our previous presentations and publications (PME 2002; PME NA 2002 and 2001) present our findings with respect to the strategies that teachers have incorporated into their classroom instruction as a result of the test. This paper addresses an issue not heretofore addressed: the issue of “teaching to the test” and how that manifests itself in the context of classroom instruction.

³ We are grateful to William Firestone and Lora Monfils who did much of the data collection and analysis for this research.

this manifests itself in classrooms. The test that we examine is the NJ Elementary School Performance Assessment (ESPA), introduced in 1999. This test was designed to be consistent with state and national standards, and measure student achievement in several content domains including mathematics. The test combines multiple choice and constructed response items, but does not involve hands-on performance assessments. Further, this test was implemented with relatively low stakes attached to it. More specifically, at the student level, test scores may be used in-house for identifying students in need of instructional support, but are not used for grade retention or promotion. Although test scores may be used in state monitoring of districts having special needs and receiving supplemental state funding, the more general effect is local competition among schools and districts motivated by the annual release of school scores in local newspapers and state websites. The teachers we interviewed, generally speaking, also reported that they perceive the test to be of low stakes. In fact, despite reservations about the test's length and the time taken away from instruction, teachers generally view the test as a positive force for improving instruction. Because the stakes are relatively low, the NJ context provides a useful study of what teachers will do in terms of "teaching to the test" even when they do not have a strong accountability system motivating them to increase their students' performance on the test.

Before continuing, it is useful to provide a review of the arguments surrounding the impact of testing on classroom instruction, especially as it relates to test preparation to provide a context for considering our analysis. To begin, many supporters of standardized testing see testing as one way of using the authority of the state to ensure that all students are exposed to the same instructional standards. They maintain that such tests can prompt teachers to revise their instructional approaches to conform to the expectations set forth by the test (and standards that they are based upon), and that sort of teaching to the test can be good if the test is well designed. (O'Day & Smith, 1993). There are others, however, that believe that extensive testing will encourage measurement of less relevant skills, and reinforce more traditional approaches to teaching (McNeil, 2000). Many contend that tests can and do influence the actual content taught in classrooms. For example, untested content either falls out of the curriculum or gets put off until the end of the year and after the test (Corbett & Wilson, 1991; McNeil, 2000). The types of items that are placed on the test are also claimed to influence the types of problems teachers use in class. The argument is that by including items that require students to solve more complex problems, teachers will be more likely to provide students with the opportunity to do the same in class. Indeed, one reason given for the great interest in various forms of performance assessment and portfolios in the 1990s was the hope that tasks requiring students to show their work and explain their answers would promote inquiry-oriented instructional approaches (Resnick & Resnick, 1992; Rothman, 1995). Many tests do indeed combine both multiple-choice formats with other formats intended to measure higher order thinking and problem solving abilities. However even when tests employ formats where students construct responses, researchers note that some of the same types of instructional reactions that are typical of the more traditional formats have been found to occur (Smith, 1996; Stecher & Barron, 1999). Regardless of the format, some researchers report that the evidence that testing promotes instructional change remains unconvincing or inconclusive at best (Newmann, Bryk, & Nagaoka, 2001; Smith, 1996).

It is therefore important to study and document the implications of testing on actual classroom practice, both in contexts where the stakes are high and low. NJ, as mentioned previously, provides a useful case study of a situation in which the stakes are relatively low.

This paper will report on some issues associated with teaching practices related to preparation for the test. The research question we address here is the following: What is the relationship between actual instruction and test preparation? In brief, we will report that an increased focus on preparation for the test occurs widely and is multidimensional ranging from embedded approaches that emphasize the use of test-like activities and items to drill on low-level mechanics.

METHODS

Data Collection and Sample

The qualitative aspect of our study involved collecting data through direct observation and interviews for a total of 78 teachers. The teachers came from districts that were representative of the state in terms of geography and demography. There were 63 teachers observed in 2000, (58 of those were observed twice, the others only once). There were 27 teachers observed in 2001, (26 of the 27 were observed twice, the other only once). Of the total of 78 teachers, 12 teachers were observed in both years.

After each observation, teachers were interviewed about such issues as the content of their lessons, the nature and goals of the instruction and the effect of the ESPA upon various aspects of their classroom practice. A researcher, who kept a timed running record of all activity that took place in the classroom, particularly noting the activity of the teacher and the students, collected the qualitative data. The field notes included problems and activities, materials used, the questions and responses of students or teachers, the physical layout of the rooms, the overall sense of classroom community, and any other elements of the class that could be observed. All observations were coded by at least two independent raters (for more information on the coding system and results, see Firestone, Schorr and Monfils, in press)

INTERVIEWS

After each observation, teachers were interviewed and asked several open-ended questions about their instructional practice. These questions were related to the lesson itself, asking about the objective of the activity, how it fits into the math curriculum and the instructional methods that were used. Teachers were asked reflective questions such as how successful they felt the lesson was and what they would change when doing this lesson in the future. They were also asked about their own professional development and about how the 4th grade test has impacted their practice in terms of curriculum and specific test preparation.

Interview data were transcribed and entered into a qualitative data analysis software package. Interviews were sorted by question. Responses were analyzed in clusters, as there was considerable overlap in responses given to individual questions. Within each cluster, responses to specific questions on test preparation practices were reviewed and coded according to emergent themes. Responses were counted within each code.

RESULTS AND DISCUSSION

Our results indicate that the teachers we observed and interviewed are indeed making changes in their instructional practices in order to help their students do better on the test. Throughout this section we will describe how this manifests itself in classrooms.

The changes that teachers were making in response to this low stakes test were in most instances, embedded into their everyday classroom practice, or more focused on specific procedures and practices in the periods just prior to the actual test⁴. Some of the more common practices that occurred in the weeks just prior to the test included drill on “test-besting” skills (ways to learn what types of questions are routinely on the test, and how the test is organized and scored in order to increase the likelihood of answering a question correctly) and the use of test-like items either for practice sessions throughout the year or as part of classroom activity during the month before the test. For example, in the year 2000 post-observation interviews, most teachers (37 out of 58) noted that they used either practice books or sample test problems they received from the State or downloaded from the Internet. Twenty specifically stated that they use commercially available ESPA prep books. The amount of time spent using such workbooks ranged from strictly the month before the exam to once a week starting in September (the beginning of the school year). Another 17 said they use their own sample problems or sample problems or tests they received from the state or downloaded from the state’s website for the same purpose.

Most of the teachers involved in our study noted that they like to use the sample questions or tests so that their students can become familiar with the format and style of exam when it is administered (in late spring). Some teachers use ESPA-like questions as a “problem of the day” throughout the year, while others only use them closer to the actual time of the exam as a means by which their students can practice a “timed” test. In our 2001 interviews, when we asked specifically about test prep in the few weeks before ESPA, an overwhelming majority (19 of the 27 teachers) responded said that they used commercially prepared materials or sample problems released by the state, another 12 stated that they taught test-taking skills. Only 10 (of the total of 27) of those questioned stated that they reviewed the curriculum as a way to prepare their students for the test.

To better understand our findings we looked to the interview data to see what teachers thought about the effects of ESPA on their actual teaching as it relates to preparing students for the test. Many teachers described a more embedded approach to test preparation that included the use of more open-ended types of questions throughout the year. For example, one teacher stated the following.

I like the way the questions are challenging and make them think. I think it certainly has affected the way that I've taught, I teach, and that I'm very, I'm always looking for opportunities to have an open-ended question somewhere and that's good.

The following quote from one teacher is typical of what teachers across the state told us.

⁴ See also Monfils, et al, 2002

The benefits of the test are that I think it is making district sit up and take notice that you know what, we can't teach the way we used to teach because the kids are not getting the skills they need for real life. ...the main thrust of the test is to change the way teachers are teaching in the classroom to match the skills that are needed for the students of today. We can't teach knowledge. We have to teach skills for getting knowledge. That was what the test was supposed to drive was this kind of teaching in the classroom. ...

Most teachers made reference to some form of an embedded approach to preparing their students for ESPA, using phrases like “teaching life skills”, “teaching that reflects the standards”, “critical thinking”, “yearlong preparation”, and stating that by covering the curriculum and corresponding standards, they were indeed preparing their students for ESPA.

My classroom is set up to work with ESPA, not separate that I'm gonna stop and train for. My writing folders, right from the beginning my math portfolios--everything is open-ended questions, rubrics are used from the beginning, language is used. It's just part of our way of life in here.

Many of the teachers who claim to use a more embedded approach to test preparation still use practice items from commercial materials, especially the open-ended items. They state that they do this to familiarize their students with the format of ESPA and/or as an end of unit review. One teacher described how she used practice problems from the “Coach” book for each content domain or cluster that appeared on the test (content domains or clusters are organized around such topics as patterns, operations on numbers, etc.) after she felt that her students had learned most of the skills as part of their regular instruction from the math text.

So what I do is I take that review part at the end of each cluster, and when I feel like the class has, you know, that we've done many of the things in that cluster, then I'll pull out that review and so they'll get used to some of the types of questions that are on the ESPA. And the biggest ones that I do are those open-ended questions...

The open-ended items were mentioned by almost all of our teachers as providing opportunities for worthwhile instruction in terms of getting their students to think more about mathematics, and explain their answers orally or in writing. These teachers tended to talk more about teaching their students about process rather than particular content just for the test. One teacher stated the following.

I don't believe in teaching to the test. But the kinds of skills -- if I can incorporate something in a lesson that they need on a test, to me that's a life skill; it's not a test skill

Lessons typical of this embedded approach to test preparation were observed in a number of classrooms. In one classroom, observed just a few weeks *after* ESPA, students worked in groups on an ESPA-like problem (used as a problem-of-the-day). The teacher said that students had been taught to use rubrics to score their own work, and it was clear throughout the lesson that the students were aware of the rubric criteria for a high score. The teacher also stated that the practice of using an ESPA problem of the day was something that she did throughout the year, and chose to continue even after the test. In fact, she even encouraged students, as a homework assignment, to create their own problem of the day.

Other teachers emphasized a more direct approach to test preparation. For instance, we observed one teacher who used practice items from a test preparation book as a review to bring closure to a math lesson on standard and metric units of measurement. In this lesson the teacher demonstrated conversion between scales of measurement with water and containers and asked the students discrete, single response questions. There was little room for student inquiry, and students were not given the opportunity to explore by working with the materials themselves. In the last 5-10 minutes of class, the teacher used multiple-choice practice items from the test prep book he had distributed earlier to ask for responses to narrow questions about the units of measurement taught in the day's lesson. In another observation, the teacher used test preparation booklets but did not discuss the answers or attempt to have all students understand the concepts behind the questions.

Teachers in districts that experienced a drop in scores often referred to a more decontextualized test preparation, one that involved use of commercial materials in the period preceding ESPA. There was a great emphasis on the use of questions which they believed to be similar in nature to test questions, and the direct drilling of students on test-besting techniques. In one district a third grade teacher said:

We got our results back from last year. Apparently, the district wasn't very happy with them.... So, ah, because of that, they ordered us books. So I guess we're going to be asked to teach ESPA stuff from an ESPA book.

The overarching belief in such situations was that the use of commercially prepared books could help to increase scores on the test. One teacher bemoaned the fact that when the scores came back (they were expected in January) the teachers would undoubtedly be forced to begin the test preparation activities. She said that the scores would come “Just in time to start test-besting again.”. Another teacher who used test like items to get the students ready for the test openly stated that she did not like the term “teaching to the test”, rather she said “and it's not that you're teaching to the test, but you're prepping for the test”.

CONCLUSIONS

The question that we have tried to address is: “What is the relationship between actual instruction and preparation for the test? Our findings suggest that teachers are indeed making choices about curriculum and practice, which are motivated solely by the test. Almost all teachers used some form of test preparation in their instruction. In some cases, this means “teaching to the test” by drilling students on test mechanics and/or simplistic notions about how to get higher scores by out-guessing the test makers. In other cases it means incorporating some embedded strategies into everyday instruction. This includes, for example, using more open-ended problem activities that resemble the test questions throughout the year, or in the period just prior to the test.

In sum, our findings reveal that even though the stakes are relatively low, the test is indeed causing teachers to consider ways to get their students to increase their scores—many of which, we do not believe contribute to learning with understanding, especially in cases where the emphasis is on “out smarting” the test makers and scorers. On a brighter side, we have noted that some teachers have incorporated more open-ended activities into their regular classroom instruction as a way to familiarize students with the

test-like questions. These attempts did not always result in increased opportunities for learning either. However they may hold greater promise as we consider ways to capitalize on teachers' good intentions given the presence of such testing.

References

- Corbett, H. D., & Wilson, B. L. (1991). *Testing, reform, and rebellion*. Norwood, NJ: Ablex.
- Firestone, W.A., Schorr, R.Y. & Monfils, L. (in press). *The ambiguity of teaching to the test*. Lawrence Erlbaum Publishers, Mahwah, NJ
- Firestone, W.A., Monfils, L., Mayrowetz, D. & Camilli, G. (2001) *The ambiguity of teaching to the test: A multi-method analysis*. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.
- McNeil, L. M. (2000). Contradictions of school reform: Educational costs of standardized testing. New York: Routledge
- Monfils, L., Firestone, W.A., Schorr, R.Y., Hicks, J., Polovsky, T. (2002). *Teaching to the test in New Jersey: A follow-up study*. Paper presented at the Annual Meeting of the American Educational Research Association. New Orleans, LA.
- National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA.
- New Jersey Mathematics Coalition and New Jersey State Department of Education. *New Jersey Mathematics Curriculum Framework*, 2002.
- Newmann, F. M., Bryk, A. S., & Nagaoka, J. K. (2001). *Authentic intellectual work: Conflict or coexistence*. Chicago: Consortium on Chicago School Research
- O'Day, J., & Smith, M. S. (1993). Systemic reform and educational opportunity. In S. H. Fuhrman (Ed.), *Designing coherent educational policy* (pp. 250-312). San Francisco: Jossey-Bass.
- Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement, and instruction* (pp. 37-75). Boston: Kluwer Academic Publishers.
- Rothman, R. (1995). *Measuring up: Standards, assessment, and school reform*. San Francisco: Jossey-Bass
- Schorr, R.Y., Bulgar, S., Razze, J., Monfils, L., Firestone, W.A. (in press). Teaching fourth grade mathematics and science. In Firestone, Schorr, Monfils (Eds.). *The Ambiguity of Teaching to the Test*. Lawrence Erlbaum, Hillsdale, NJ.
- Schorr, R.Y. & Bulgar, S., Monfils, L. & Firestone, W.A. (2002). Teaching and testing at the fourth grade level: An analysis of a two-year study in New Jersey. *Proceedings of the North American Chapter of the International Group for the Psychology of Mathematics Education XXIV*. Athens, GA.
- Schorr, R.Y. & Firestone, W.A. (2001). Changing mathematics teaching in response to a state testing program: A fine-grained analysis. Paper presented at the Annual meeting of the American Educational Research Association, Seattle, WA.

- Smith, M. L. (1996). Reforming schools by reforming assessment: Consequences of the Arizona Student Assessment Program. Tempe, AZ: Southwest Educational Policy Studies, Arizona State University.
- Stecher, B. M., & Barron, S. I. (1999). *Quadrennial milepost accountability testing in Kentucky*. Los Angeles: CRESST