Education Policy Analysis Archives

Volume 8 Number 35 **July 26, 2000** **ISSN 1068-2341**

A peer-reviewed scholarly electronic journal
Editor: Gene V Glass, College of Education
Arizona State University

Articles appearing in **EPAA** are abstracted in the *Current Index to Journals in Education* by the ERIC Clearinghouse on Assessment and Evaluation and are permanently archived in *Resources in Education*.

State Standards, Socio-fiscal Context and Opportunity to Learn in New Jersey

William A. Firestone
Gregory Camilli
Michele Yurecko
Lora Monfils
David Mayrowetz

Rutgers University

Abstract
A survey of 245 New Jersey teachers provides a baseline for examining how the introduction of state standards and assessments affects the teaching of math and science in the 4th grade. These policies are promoting teaching of additional topics in both areas. The changes in the delivery of professional development have not yet been sufficient to lead to substantial changes in instructional practice. While inequities in access to material that characterized the state in the early 1990s have diminished, we find a pattern of inquiry-oriented science teaching more prevalent in wealthy districts and teaching to the test more prevalent in poorer ones. We also note some areas where middle-income districts appear disadvantaged.

A central goal of the standards movement has been to help all children learn challenging content (Smith & O'Day, 1991). Forty-four states have now adopted standards for student proficiency in the core academic areas, 41 states have aligned assessment with their math standards, and 25 have aligned assessment with their science standards (Quality Counts, 2000). While great attention is being paid to what students are learning, less scrutiny has been given to what they are taught. Yet, the former depends at least in part on the latter (Wiley & Yoon, 1995). For that reason, state standards are intended to provide guidance on what should be taught, as well as what students should learn (Smith, Fuhrman & O'Day, 1994).

The adoption of standards and assessments does not guarantee students access to instruction, especially for poor students. For that reason, people have begun to worry more about "opportunity to learn" (OTL) or "whether or not… students have had an opportunity to study a particular topic or learn how to solve a particular type of problem presented by a test" (Husen as cited in McDonnell, 1995, p. 306). Advocates for minorities have seen the reporting of OTL standards as a way of ensuring that poor and minority students are not disadvantaged inappropriately when standards are raised. As one observer noted, without OTL standards, "you don't know if the school if failing, or if students are failing" when test scores are low (Rothman, 1993, p. 21).

Both the federal and state governments have been much more willing to adopt student performance standards than OTL standards since the latter specify the government's obligation to deliver services to students (McDonnell, 1995). Moreover, the legal mandate for guaranteeing that OTL be provided is ambiguous, even though the issue arose in the early years of state testing. According to Millman and Green (1989, p. 356):

> The court decision in the *Debra P. vs. Turlington* (1981) case seems to have established the necessity that, at least for certification tests for high school graduation, the tested material must consist of content that is currently taught, that is, the student must have been provided adequate preparation and, thus, had a fair opportunity to learn the material.

Precise requirements of a fair opportunity to learn remain ambiguous.

Several decades of research have indicated how difficult it is to change teaching practice (McLaughlin, 1990; Cuban, 1993). Simply imposing standards by decree is not likely to modify teaching practice if teachers do not understand what is expected of them or have the resources to carry out a standards-based program of instruction. The situation can be especially challenging in mathematics and the sciences where elementary education teachers may lack the background knowledge to effectively teach more challenging content.

This article introduces a project designed to explore how state standards and related policies influence teaching practice. In May, 1996, New Jersey announced a new set of "core curriculum content standards" (NJSDE, 1996). These standards began to take practical reality for elementary school teachers when state assessments aligned with these standards were introduced in 1998. In the Spring of 1999, as the state administered its new fourth grade mathematics and science assessments for the second time (the first time for which results would actually be released publicly), we began a three-year study to examine how teachers in those grades teach mathematics and science. Using a state-wide representative survey, this article describes three dimensions of teaching practice: the content taught, access to and use of materials, and teaching to the test. In each area, we investigate what in being taught and how equitably practices are distributed among wealthy and poor districts. We also explore teachers' background knowledge and opportunities to learn about new practices. Our preliminary conclusions are that:

- The introduction of standards and assessments is broadening the range of topics taught in mathematics and science.
- A useful baseline measure for assessing teaching to the test can be developed.
- Opportunities remain limited for elementary teachers to learn the new knowledge required to improve their mathematics and science teaching.
- The inequities between wealthy and poor districts are complex and may be overstated, but there is clearly more teaching to the test in poor, urban districts and more hands-on science teaching in wealthier districts.

Before addressing these issues we describe the context for standards implementation in New Jersey and the research methods employed in the study.

The Policy Context

In the last decade educational policy in New Jersey has been driven by two related phenomena: school finance litigation and the development of standards and related assessments. Whereas financial resources can influence the distribution of OTL, legal battles surrounding the school finance issue also motivated the adoption of standards.

School Finance Litigation

Since school finance litigation began in New Jersey thirty years ago, there have been two court cases, eleven decisions, numerous school finance bills, and other laws and regulations (Goertz & Malik, 1999). The litigation and related legislation has focused on whether the state was obligated to provide all children therein a "thorough and efficient education." While these actions have had a number of implications for education in New Jersey, two are especially critical here: the definition of a thorough and efficient education, and the financial provisions to ensure that all children could receive one.

The court has been reluctant to define a thorough and efficient education except in the broadest terms:

> For those special needs districts [the approximately 30 poor urban districts identified by the court as inequitably served by the state], a thorough and efficient education—one that will enable their students to function effectively in the same society with their richer peers both as citizens and as competitors in the labor market—is an education that is the substantial equivalent of that afforded in the richer districts (*Abbott v. Burke*, 643 A.2d 575, 580 (1994)) (Abbott III)

Beyond stating that children in poor districts should get the same education as those in wealthy districts, this decision provided very little guidance; and the court continued its multi-year effort to urge the state department of education to specify criteria in more detail. This was accomplished in part in the Comprehensive Plan for Educational

Improvement and Financing (CEIFA), the school funding law of 1996, which defined a thorough education as one in which children succeeded in meeting the 56 outcomes specified in the Core Curriculum Content Standards. Thus, the standards became the criteria for educational effectiveness, and state tests administered in 4th, 8th, and 11th grade would operationalize those criteria. The court found that these standards and assessments were "the first real effort on the part of the legislative and executive branches to define and implement the educational opportunity required by the Constitution… and are facially adequate as a reasonable legislative definition of a thorough and efficient education" [*Abbott v. Burke*, 693A.2d 417, 428 (1997) (Abbott IV)].

This effort was not sufficient to clarify what constituted adequate educational funding for all children in the state. Thus, the court continued to use a two-part yardstick. First, the poorest districts in the state should spend essentially the same per capita as the wealthiest districts (Goertz & Malik, 1999). The state had developed a classification of districts (District Factor Group or DFG) based on a composite measure of community, social, and economic variables such as the educational and occupational background of the population, per-capita income of the district, and mobility. The DFGs were designated by letter with the poorest districts labeled "A" and the wealthiest labeled "J". Per-pupil spending in the special needs districts designated by the court was expected to match that of the highest DFG districts. As late as 1993-94, the 14% of districts were spending 22% more than the poorest although their collective tax rate was 43% lower (Firestone, Goertz & Natriello, 1997).

Second, in addition to equal base spending, the court required the state to support a series of supplemental programs for the poor urban, districts. Urban schools were expected to implement a whole school reform program model such as Success for All (Porter, 1999), extend early childhood education services to 3- and 4-year olds, and began programs to refurbish aging and decaying buildings. Since these programs could not be supported locally, they had to be underwritten by the state (Goertz & Malik,1999; Erlichson, Goertz, & Turnbull, 1999). By the 1999-2000 school year, the equal base funding provisions were in place and implementation of the special programs had begun although not without disputes about the local level of funding and district discretion in designing their whole-school reform and early childhood programs.

Equal basic funding is an important development, and extremely unusual in a state noted for inequities in education. In 1996 only two states had a greater dollar gap in spending between the fifth and 95th percentile districts than New Jersey (Quality Counts, 2000). However, the court remedies and new funding formula did not extend to all districts. Schools in DFGs as low as B and into the middle of the fiscal distribution were spending less per child than either the wealthiest or the poorest districts in the state.

Standards and Assessments

As a normative perspective, standards theory recommends that state standards become the criteria with which assessments are aligned. However, like many American states, New Jersey began with assessments rather than standards. Its first testing system, begun in the late 1970s, was designed to measure "minimum basic skills" as a means of maintaining the accountability of poor urban districts, who at that point were receiving a new infusion of state funds. Several revisions ensued, and by the early '90s the keystone of the state's testing system was the High School Proficiency Test (HSPT), administered in 11 th grade as a partial requirement for high school graduation. This test covered mathematics, reading, and writing at a more challenging level than the earliest test, but passing score was still set at a basic skills level. The HSPT was accompanied by an

Early Warning Test (EWT), given in 8th grade to help schools identify children at risk of failing the graduation test. These tests had special significance to educators because patterns of low scores on these tests could become grounds for state takeover of a district. Districts were also expected to administer conventional achievement tests of their own choice at grades not tested by the state (Firestone et al., 1997).

During the 1990s as the standards movement took hold nationally, teams of content experts and teachers were formed within the state to write the core curriculum content standards in seven curricular areas as well as a set of cross-content workplace readiness standards. These efforts were heavily influenced by national standards documents in mathematics and science and became official in May, 1996 (NJSDE, 1996). The resulting standards for mathematics and science are listed in Appendix A. These core standards are accompanied by cumulative progress indicators for grades 4, 8, and 12. Separate documents provide curriculum frameworks to offer guidance to educators in implementing the standards.

The state is now phasing in 4th, 8th, and 11th grade tests that are intended to be aligned with the standards in each area. The degree of alignment to the standards is difficult to assess because—as in many states—strict confidentiality is maintained over operational test items. This creates difficulties for educators who wish to be given test results item by item in order to seek an easier method for aligning their instruction more closely with the assessments.

The current tests are an effort to move away from the basic skills or advanced basic skills orientation that characterized earlier state tests. The 4th grade mathematics tests include 32 closed-ended and five open-ended items; and the matrix for selecting items includes a dimension of "problem-solving skills" with categories like "procedural knowledge, conceptual understanding, and problem-solving skills" (NJSDE,1998, p. 6). The 4th grade science test is similarly organized. One sample open- ended item and one sample closed-ended item from the test specifications are included in Appendix A. The 4th grade mathematics and science tests were first administered in the spring of 1998, but because of technical problems scores were not released. The following year scores were released in the fall after the spring 1999 administration.

The introduction of new standards and assessments in mathematics and science should provide clarity regarding what is expected to be taught in each area, and ensure that these subjects receive consistent attention. Whether this attention takes the form of short-term "teaching to the test" or deeper changes in practice, and whether access to new forms of instruction is equally distributed in the state remains to be seen. Recent court and legislative actions may further stimulate access to new forms of instruction. We turn now to the survey designed to address these issues.

Study Sample

In the spring of 1999, we initiated a three-year study to examine teachers' response to the new testing program in the areas of mathematics and science. Data were collected from a statewide sample of 4th grade teachers. Just over 600 teachers were asked to respond to a complex set of instruments. After extensive telephone follow-ups and remailings, 245 teachers completed a telephone survey, 172 completed an additional mailed questionnaire, and 110 provided examples of mathematics and science lessons they taught, including materials given to students and more detailed reports on teacher and student activities conducted with those materials. (Note 1) The sample is highly representative with regard to district wealth as measured by DFG (See Table 1).

Past research suggests that successful change in teaching practice depends on opportunities for teachers to learn new practices required by the policy (Cohen & Barnes, 1993; Firestone et al., 1998). However, the kind of professional development that is most likely to lead to substantial change in practice continues to be rare (Loucks-Horsley, Hewson, Love, & Stiles 1997). In order to assess the effects of professional development, we sought to oversample schools that were known to engage in extensive professional development with respect to mathematics and science. The New Jersey State Systemic Initiative shared with us results of a survey identifying districts engaged in the most extensive professional development in those subjects. We attempted to ensure that 25% of our sample came from these districts. In fact 49 of the completed telephone interviews (20%) and 30 of the completed mailed questionnaires (17%) came from high professional development districts.

Table 1
Distribution of Responses by DFG

District Factor Group							
	AB: (Poorest)	CD	DE	FG	GH	IJ: (Wealthiest)	Total
Interviews	71	29	32	24	35	54	245
Percent	29%	12%	13%	10%	14%	22%	100%
Questionnaires	49	21	23	14	25	40	172
Percent	28%	12%	13%	8%	15%	23%	100%
4th Grade Students in State (%)	30%	9%	15%	13%	13%	19%	100%

In the following section we explore what content is being taught, teachers' access to materials, the extent of teaching to the test, self-reported knowledge about standards, and teachers' access to professional development.

Content Coverage

Standards and assessments are supposed to be able to influence the content taught to children. Smith (1991) and Corbett & and Wilson (1991) found that the introduction of minimum competency tests narrowed the range of subjects taught in a school to what was on the test. Firestone, Mayrowetz, & Fairman (1998) suggested that the introduction of more complex performance assessments can affect the presence and order of topics taught. There is reason to believe that the new standards and assessments are affecting content coverage in New Jersey. Fifteen percent of our sample said they were teaching more math and 14% said they were teaching more science. Noticeable changes are being made within each content area but these are different in mathematics and science.

Math Content

Traditionally, elementary mathematics has focused on basic arithmetic—addition and subtraction of whole numbers with some introduction of fractions and decimals and geometric shapes. New Jersey's Core Curriculum Content Standards expect the introduction of a wide range of content at the fourth grade level, including a broader range of geometric issues; the foundations of algebra; better understanding of measurement; an introduction to statistics, probability, and data analysis; and discrete mathematics (NJSDE, 1996). We wanted to access how teachers were using their time in mathematics and how that time use was changing. In order to avoid influencing respondents familiar with the standards terminology, we identified 17 topics that represented a mix of classic elements of the elementary mathematics curriculum and areas that were not likely to have been taught before the standards were introduced [Appendix C]. We then asked teachers how many lessons they taught each of the 17 topics, and whether they had increased or decreased the time allocated to each topic in the last three years—i.e., when the standards were being introduced and the ESPA was being given for initially.

Although we do not have a firm fix on how time was allocated to topics before the standards were introduced, it appears that the gap between conventional and newer topics is being reduced with teachers adding time to newer topics. Working with experts familiar with math teaching in the state, we identified three traditional topics: paper and pencil mathematical operations with whole numbers, adding and subtracting decimals via paper and pencil, and place value relationships (whole numbers, decimals); and three newer topics: open sentences, use of variables (strategies used to prepare students for algebra), probability, and dealing with data (collecting, organizing, analyzing, and displaying data). Most teachers reported that they spent many lessons on whole number operations: 96% spent eleven or more lessons a year on that topic. In addition, 58% devoted eleven or more lessons to place value relationships, and 22% spent that much time on adding and subtracting decimals. Although fewer teachers devoted substantial time to the newer topics, 50% spent 11 or more lessons on dealing with data. Thirty three percent spent 11 or more lessons on open sentences, and 14% on probability.

Although the larger balance of teaching time was spent on older topics, most teachers reported *increasing* the amount of time the spent on the new topics (Figure 1). In general time spent on the older topics remained fairly constant, with the exception of whole number operations. A large portion of teachers (29%) reported decreasing time spent on whole number operations. Based on this evidence, it appears that newer topics are taking a more prominent place in the curriculum, but not necessarily replacing older topics.

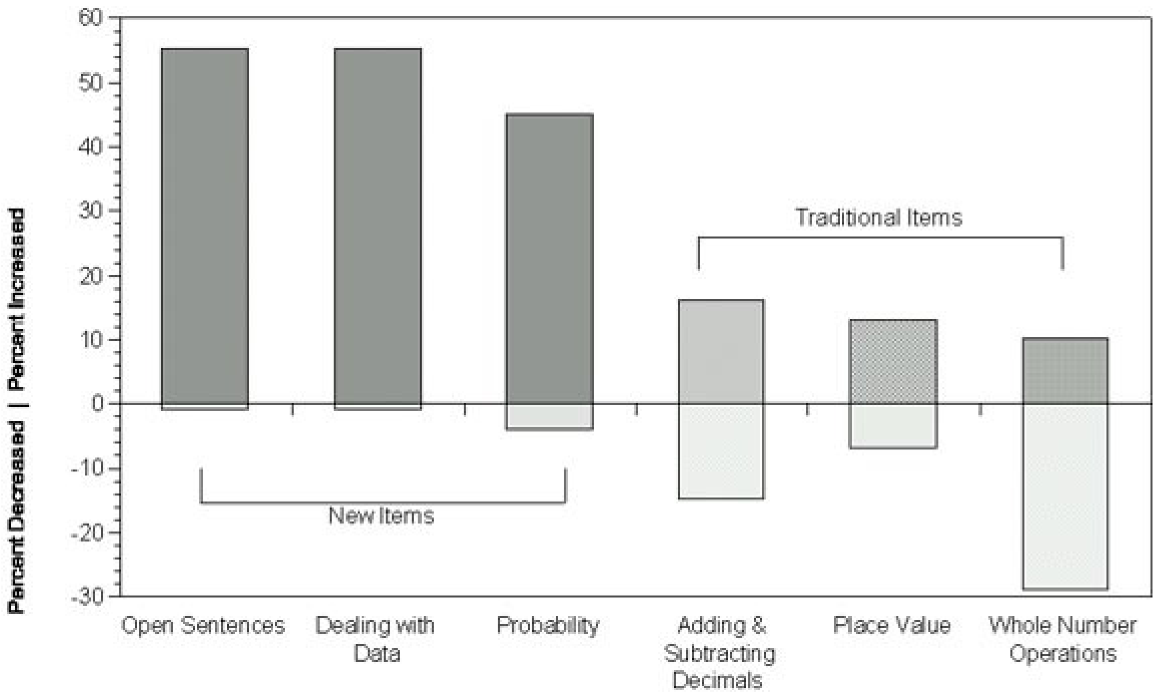


Figure 1. Percent Changes in Mathematics Items

We also explored whether the time allocated to topics was the same in wealthy and poor school districts. In 13 of the 17 topic areas there were no significant differences between DFGs. However, in four topics identified as new by our mathematics experts, we noted an interesting u-shaped pattern. Teachers in poor, urban districts and the wealthy districts spent more time on these topics than middle income districts (Table 2). An explanation for this pattern has not yet been found.

Table 2
Differences by DFG in Lessons Allocated to Math Topics
(Percent of teachers devoting 11 or more lessons to a topic, n = 151-154)

	District Factor Group			
	Abbott*	C-E	F-H	IJ
Probability	27%	12%	3%	19%
Patterns, functions	49%	16%	21%	36%
Open sentences	46%	29%	19%	41%
Discrete math	54%	25%	16%	36%

* District wealth is generally measured by DFG. The Abbott districts are all DFG A or B and have been designated by the state Supreme Court as those where spending must be equalized with wealthy districts in the state. The DFG metric runs from A (districts with large numbers of poor and generally at-risk children) to IJ with large numbers of children from wealthy families. Teachers from DFG-B districts that are not "Abbott districts" have been excluded from this comparison.

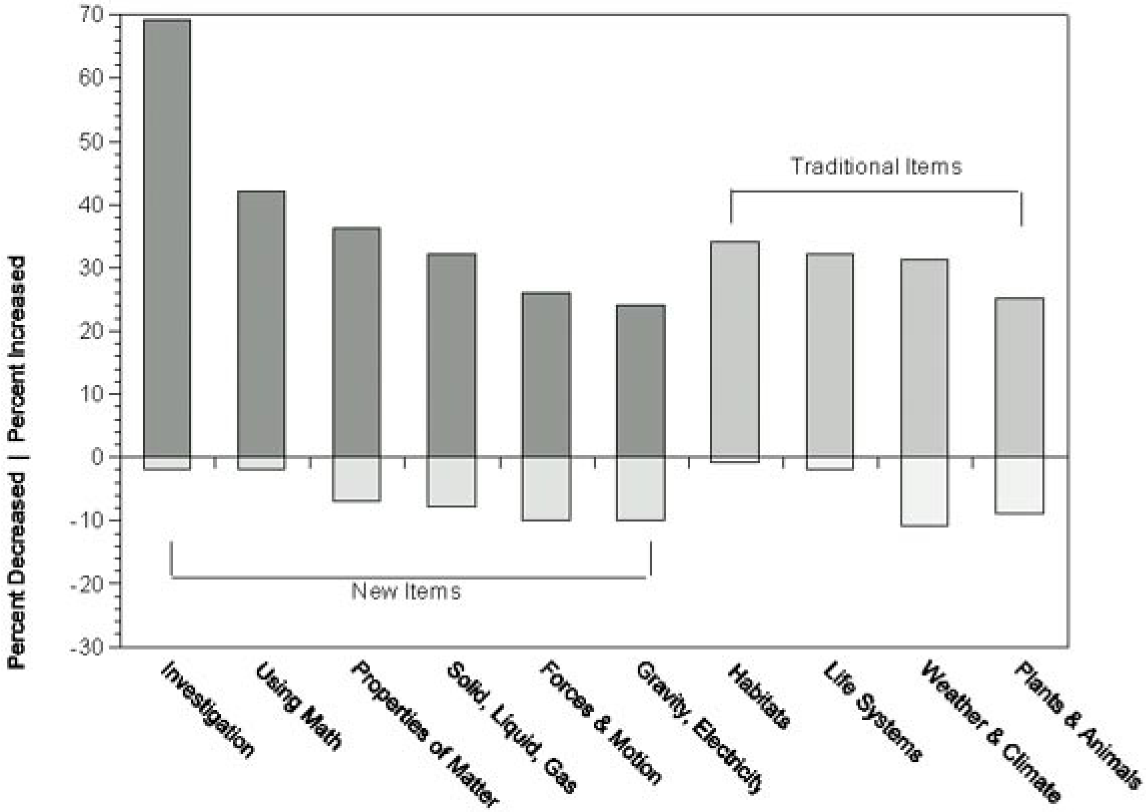


Figure 2. Percent Changes in Science Items

Access to Materials

New Jersey's Core Curriculum Content Standards place an increased emphasis on a more active role for students to take in learning mathematics and science. The mathematics standards require students to "develop an ability to pose and solve mathematical problems,… develop reasoning ability and… become self reliant independent mathematical thinkers; [and] regularly and routinely use calculators, computers, manipulatives, and other mathematical tools to enhance mathematical thinking, understanding, and power" (New Jersey State Department of Education, 1996, p. 4-9). The science standards require that students "develop problem-solving, decision-making, and inquiry skills, reflected by formulating usable questions and hypotheses, planning experiments, conducting systematic observations, interpreting and analyzing data, drawing conclusions and communicating results" (New Jersey State Department of Education, 1996, p. 5-3). These changes are in keeping with national standards which require more problem solving in mathematics and hands-on inquiry in science. At the same time they place greater demands on districts to provide additional materials—mathematical manipulatives, calculators and computers, the wherewithal for scientific experiments—beyond the basic textbooks that have been so typical of American teaching (Cuban, 1993). In fact, some textbooks include alternatives like science kits or math manipulatives.

Access to teaching equipment and supplies has historically been unequal, favoring wealthy districts. In the early 1990s, teachers in poor, urban districts reported less access to both textbooks and computers than their peers in wealthy districts. For a period of time following the passage of the Quality Education Act (QEA) which increased funding to urban districts for a short time in the early 1990s, there was some indication that poor districts were working hard to bridge the gap between themselves and wealthier districts. However, they have not been successful (Firestone et al., 1997).

The current study indicates that access to materials may be improving in poor districts. Across DFGs teachers reported having enough materials for most purposes, especially for teaching mathematics. Ninety-five percent of the teachers surveyed reported having enough math textbooks for every child to have one. (Note 2) Ninety-four percent reported having enough manipulatives for children to share, and 97% reported enough calculators for every child. The situation is nearly as good in science where 77% of the teachers reported having enough textbooks for every child, 76% reported enough science kits either for every child or for children to share, and 85% reported enough measurement and observation tools to share.

Use tends to lag behind access. Seventy eight percent of teachers report using their math texts almost every day, (Note 3) 66% use manipulatives once or twice a week, and 53% use calculators once or twice a week. The pattern in science is somewhat different. While 36% report using a textbook everyday, 40% report using it once or twice a week. Sixty-five percent report using science kits at least once a week, and 38% report using measurement and observation tools that often.

We did not identify any inequities in access to mathematics materials, supported by the high percentage of teachers who reported having enough math textbooks for every child. The situation in science is more complicated because teachers in poor, urban districts appear to emphasize the use of textbooks, while those in the wealthier districts balance textbooks with the use of science kits and other materials (Figure 3). Almost all the teachers in the Abbott districts and mid-wealth districts say they have enough science textbooks for every child and more than four fifths use them weekly. However, less than half the teachers in the wealthy districts have enough textbooks for every child and use them weekly. A third of the teachers in wealthy districts have enough kits for every child and two thirds use them weekly.

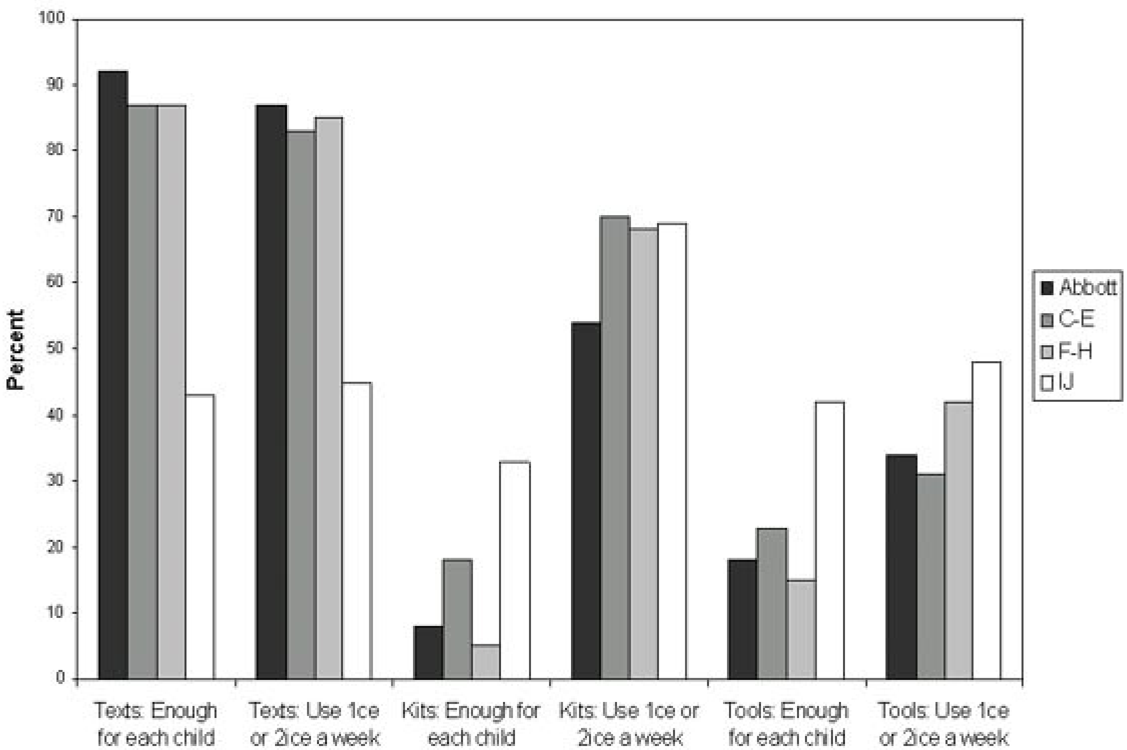


Figure 3. Access To and Use Of Science Materials

Kits are much less accessible in the poor and mid-wealth districts. Still about half the teachers in urban districts report using them weekly and use in the mid-wealth districts is comparable to that in the wealthy districts. The pattern of access to tools for observation and measurement parallels that to access to kits with substantially more teachers reporting having enough for every child in the wealthiest districts. There is a gradual trend of increasing use as one moves from the Abbott to the wealthiest districts. The reasons for these differences are not clear. However, the fact that most teachers in the state report little change in their access to materials suggests that this pattern reflects a difference in philosophy about how to teach science more than recent changes in funding.

Teaching to the Test

One of the greatest concerns with standards- and assessment-based reform has been that this strategy might lead to teaching to the test and its concomitant negative effects such as narrowing the curriculum; constricting instruction time; increasing the amount of drill while undermining efforts to promote higher order thinking skills; and increasing stress for teachers and students (Corbett & Wilson, 1991; Smith, 1991). There is also a fear that teaching to the test will undermine the validity of test results by artificially inflating test scores (Mehrens, 1998). There has been some question about whether these are inevitable effects of high- stakes accountability-oriented tests. Some have suggested that changes in test format should include more performance- oriented items and test items assessing more than mere retention of facts and computation skills might lead to tests worth teaching to and encourage teaching that promoted more conjecture, exploration, and active participation in learning (Baron & Wolf, 1996; Rothman, 1995).

To explore the distribution of teaching to the test in the state, we developed a seven-item scale with a mixture of items that seemed to reflect some of the feared negative effects of this practice and others construed as positive. The scale had an alpha coefficient of .71. Specific items included:

1. Teach test staking mechanics like filling in bubbles, how to put your name on the test, or how to pace yourself during the test.
2. Motivate students to make their best effort on the ESPA, such as suggesting they prepare by getting a good night's sleep or encouraging them to try hard.
3. Have students use rubrics to grade each other's work.
4. Teach the regular curriculum using performance-based exercises similar to the ESPA.
5. Teach test-besting skills like methods for turning story problems into arithmetic calculations or how much to write after an open-ended math item.
6. Use commercial test-preparation materials like "Scoring High" and "Measuring Up on the ESPA."
7. Give practice tests with items similar to those on the ESPA.

We asked teachers how often they performed these activities (on a scale of 1-4) all year long and the month before the ESPA was given. (Note 4) Figure 4 shows two patterns in teachers' reported teaching to the test. First, as might be expected, there is a small increase in activity during the month before the test compared to the entire year (scale mean of 2.50 for the whole year versus 2.86 for the month before the test). Second, there is a distinct pattern of teachers in the Abbott districts reporting more teaching to the test than teachers in the wealthiest districts. Teachers in the mid- wealth districts fell somewhere in between. Thus, the emphasis on test preparation as a separate activity were concentrated in the districts that most need help in improving student learning.

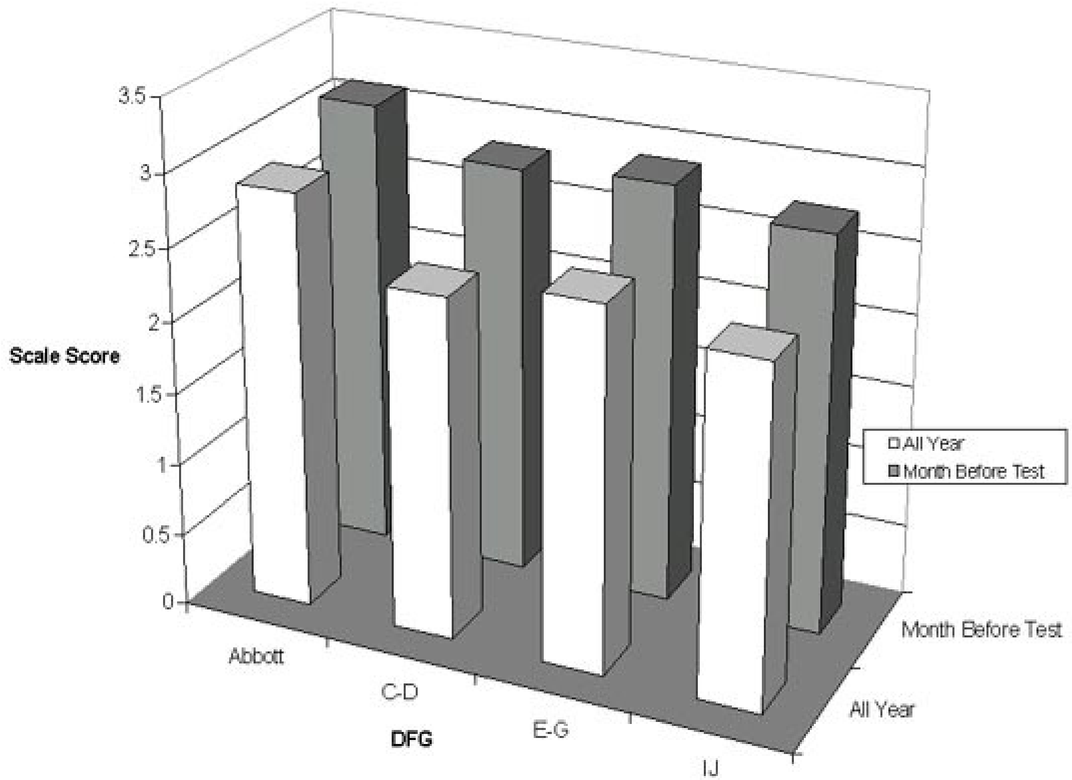


Figure 4. Teaching to the Test

Familiarity with Standards

We asked teachers to report how familiar they are with state and national standards in mathematics and science. Teachers' familiarity with state standards could contribute to changes in content taught, although central office staff who understand state standards and assessments can unilaterally change district curriculum. The national standards movement in science, and especially in mathematics precedes New Jersey's efforts by several years; and some districts were using those national standards to guide changes before state standards were adopted or tests were implemented.

Teachers were much more familiar with state than national standards. Fifty-seven percent said they understood the state's mathematics standards well, (Note 5) and 53% say they are understand the science standards well. In contrast, only 28% said that they understood the national mathematics standards well and 16 said they understood the national science standards well. Even if teachers overestimated their understanding of the standards, the state's effort has increased attention to standards-based teaching here.

For the most part, understanding of standards is equally distributed across wealthy and poor districts. The one exception is the national mathematics standards where there is a complicated pattern of differences between districts (Table 3). Generally, more teachers in the wealthy districts believed that they understandd the national standards well. However, it is not true that most teachers in the Abbott districts have limited familiarity with the national math standards. The largest concentration having moderate familiarity is in the Abbott districts while the almost two thirds of the CE teachers have only limited familiarity with the national standards. One possibility is that the wealthy districts have sought to adopt the national standards for a long time. Growing familiarity in the Abbott districts may reflect a mix of three factors: a side effect of the attention to standards in general from the adoption of state standards, the special pressures placed on the Abbott districts by the state as a by-product of the series of court cases and large amount of state money going to those districts (Firestone & Nagle, 1995), and the additional funds coming from CEIFA after the Abbott IV decision.

Table 3
Understanding of National Mathematics Standards by DFG
(Percent of Teachers, n = 158)

	District Factor Group			
	Abbott	C-E	F-H	IJ
Limited*	37%	63%	32%	33%
Moderate**	47%	21%	29%	28%
Extensive***	16%	16%	40%	39%

* Awareness only and read through once or twice.
** Understand somewhat (can implement parts in class)
*** Understand well (can implement fully in class) and expert (could lead workshop)

Professional Development

Past research on policy implementation in a variety of fields suggests that regardless of changes in incentives and punishments, teachers will not change their practice until they have learned how to perform the new tasks expected of them (Berman 1986, Cohen and Barnes, 1993). Firestone and colleagues (1998) suggest that one reason state-administered performance-based assessment has had limited impact on teaching is because teachers have had limited opportunities to learn the new content and pedagogy required by the new assessments.

Teachers reported on several dimensions of their professional development experience. Regarding the source of professional development, most learning opportunities for teachers came directly from the district. Sixty seven percent of teachers reported that some time in their district-provided professional development days in the last year had been devoted to mathematics or science. In the last year, 40% had mentored student teachers or first year teachers, 41% had served on district curriculum development or textbook selection committees, and 21% had served as lead or specialist teachers helping other experienced teachers in their district. All of these are learning experiences even though they may involve helping others.

Relatively few teachers had opportunities to develop new knowledge by interacting with experts from outside the district. Eighteen percent had taken a college course in math, science, or math or science education in the last year. Twenty two percent had participated in one the programs for improving math and science teaching supported by the National Science Foundation through its State and Local Systemic Initiatives or the US Department of Education through its Eisenhower grants to institutions of higher education. Given elementary teacher's reputation for aversion to mathematics and science, these numbers are fairly reasonable. However, since the objective is to achieve statewide high quality mathematics and science teaching, it seems quite unlikely that teachers' understanding of effective practice will grow quickly unless more avail themselves of these opportunities.

One recurring criticism of professional development is that it is usually provided through one-shot workshops where teachers receive limited and often inapplicable information with little or no follow up to help in using what they are supposed to have learned. That seems to have been the case among New Jersey's fourth grade teachers (Table 4). Only about one fifth of the teachers reported having more than two days of professional development on either content and instruction in science and math. Slightly fewer received more than two days of professional development on strategies to help students score high in math or science. It is somewhat encouraging that teachers received about as much professional development on the underlying content and instructional issues as they did on strategies to raise test scores. On the other hand, only one in 20 received more than two days on using assessment results. It is particularly disconcerting that teachers received so little support in using assessment results to improve instruction, although this may be because the state had not yet reported any ESPA results to schools when this survey was conducted.

Not only is professional development limited, so is follow up. Between 20% and 30% of the teachers report being visited later by a trainer. Follow up by principals is more common, but principals are often less well informed about the content of professional development. Their follow up may show concern and signal that the material covered is important, but substantive assistance is likely to be less than that coming from an expert. Nevertheless, between one third and one half the teachers found the professional development they received to be very useful. This may be in part a reflection of the growing demand for help in this area.

Table 4
Time in Professional Development
(Percent Reporting Various Categories)

	More than 2 days PD in year	Follow-up by trainer	Follow-up by principal	PD is very useful
Content and instruction in science	22%	25%	22%	44%
Content and instruction in math	20%	25%	26%	48%
Using assessment results	6%	21%	35%	30%
Strategies to score high in math	19%	29%	33%	48%
Strategies to score high in science	14%	22%	29%	41%

Where New Jersey teachers received more professional development, they found it more useful. The correlation between the amount of time spent in professional development and its perceived utility were .66 for content and instruction in science, .63 for content and instruction in mathematics, and .61 for using assessment results. They were lower for strategies for scoring high in math and science (.44 and .40, respectively). These findings suggest that extensive professional development efforts will be most helpful when helping teachers better understand the underlying material in a subject and effective strategies for helping students learn it. Longer time investments may also pay off for helping teachers to use assessment strategies to improve practice. Comparable concentrations are probably not as necessary to give teachers strategies to raise test scores.

Discussion

While there are limitations to what can be learned about changes in teaching practice from one administration of a survey that focuses on elementary school mathematics and science, the data presented here suggest some tentative conclusions and raise questions about two issues: ongoing changes in practice, and differences between wealthy and poor districts.

Statewide, it appears that the topics taught as part of the 4th grade curriculum are changing. This may have implications for elementary curriculum in general. In mathematics, what had been an unremitting diet of whole number facts is being leavened with other topics like probability and dealing with data. Generally, more science is being taught, and the small sampling of biology and meteorology is being expanded. There is a large increase in attention to the process of scientific investigation, some increase in attention to the introduction of chemistry and at least a smattering of attention to physics-related topics. These changes help prepare children to use mathematics as part of their adult life and give them an introduction to a broader range of science topics.

The simple addition of topics may be a mixed blessing, however. One criticism of mathematics teaching in the past has been that too many topics are taught at too little depth (Schmidt, McNight, & Raizen, 1997). The addition of new topics to the state standards could exacerbate such shallow coverage. The quality and depth of coverage is difficult to assess with surveys; hopefully, direct observation in classrooms, which is currently underway, will help address this issue. It will also be useful to collect longitudinal data on coverage of content areas to verify that the changes we believe are happening are in fact taking place. Teachers are also becoming more familiar with the state standards, and believe they are more familiar with state than with national standards. We suspect that the extent of their familiarity is overstated. Again, we hope to learn more from direct observation.

On the equity front, the picture is mixed. The good news is that some of the obvious inequities in access to materials that were prevalent at the beginning of the decade appear to be fading. However, there are hints that two pedagogies may be developing in the state: one for children in districts serving the poor, and another for districts serving the wealthy. Pedagogy in the poor districts may come to be dominated by conventional, textbook-oriented teaching and teaching to the test, while wealthier districts seem to be moving towards more exploratory, active modes of learning that are less dependent on textbooks and less driven by state tests. If so, the reasons are likely to have less to do with differences in funding and more with heavier pressures to comply with state expectations in urban districts and the challenges that come with teaching poorer children (Natriello, Pallas & McDill, 1990).

There is also the issue of those districts in the middle of the DFG distribution. These more working-class districts are not as well funded as either the Abbott districts or the wealthy districts. There are some indications that teachers in the Abbott districts are moving faster than those in the poorer of the mid-wealth districts to embrace the standards and introduce new topics to the curriculum. How strong this trend is, whether it will continue, and what its implications are for teaching practice and student achievement remain to be explored through further surveys and direct observation in classrooms.

Notes

This article was presented as a paper at the Annual Meeting of the American Educational Research Association in New Orleans, LA, April, 2000. We wish to thank Warren Crown, Roberta Schorr, John Shafransky, Sharon Sherman, and Carol Stearns for their assistance. This research was supported by a grant from the National Science Foundation. The opinions expressed are those of the authors. Neither the Foundation nor Rutgers University is responsible for them.

1. The teacher work samples are not used in this report.
2. The choices offered teachers were none, one or two to demonstrate in class, enough for children to share, and enough for every child to have one.
3. The options were almost every day, once or twice a week, once or twice a month, once or twice a semester, and never.
4. Respondents were asked to report on a 4-point scale where 1 was "almost never" and 4 was "almost always."
5. The actual choices were "Awareness only, read through once or twice, understand somewhat (can implement parts in class), understand well (can implement fully in class), and expert (could lead workshop)." The responses reported are for the last two combined.

References

Baron, J.B. & Wolf, D.P. (1996). *Performance-based student assessment: challenges and possibilities*. Chicago, IL: University of Chicago.

Berman, P.E. (1986). From compliance to learning: Implementing legally-induced reform. In D. Kirp & D. Jensen (Eds.), *School days, rule days*. (pp. 46-62). Philadelphia, PA: Falmer.

Cohen, D.K., and Barnes, C.A. (1993). Pedagogy and policy. in D.K. Cohen, M.W. McLaughlin, & J.E. Talbert (Eds.) *Teaching for understanding*. (pp. 207-239) San Francisco: Jossey Bass.

Corbett, H.D. & Wilson, B.L. (1991). *Testing, reform, and rebellion.* Norwood, NJ: Ablex.

Cuban, L. (1993). *How teachers taught: Constancy and change in American classrooms, 1890-1980*. (2nd ed.) New York: Teachers College Press.

Erlichson, B.A., Goertz, M. & Turnbull, B. J. (1999) *Implementing whole school reform in New Jersey: year one in the first cohort schools*. New Brunswick, NJ: Rutgers, the State University of New Jersey.

Firestone, W. A., & Nagle, B. (1995). Differential regulation: Clever customization or unequal interference? *Educational and Evaluation and Policy Analysis*, 17(1), 97-112.

Firestone, W.A., Goertz, M.E., & Natriello, G.J. (1997) *From cashbox to classroom: The struggle for fiscal reform and educational change in New Jersey*. New York: Teachers College Press.

Firestone, W.A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, *20*(2)95-113.

Goertz, M. & Malik, E. (1999) In search of excellence for all: the courts and New Jersey school finance reform. *Journal of Education Finance, 25*(Summer)5- 32.

Loucks-Horsley, S., Hewson, P.W., Love, N., & Stiles, K.E. (1997) *Designing professional development for teachers of mathematics and science*. Thousand Oaks, CA: Corwin Press.

McDonnell, L. M. (1995). Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis*, *17*(3), 305-22.

Mehrens, W. A. (1998). Consequences of Assessment: What is Evidence? [Online]. *Educational Policy Analysis Archives*, *6*(13). Available: http://epaa.asu.edu/epaa/v6n13.html

Millman, J. & Green J. (1989). The specification and development of tests of achievement and ability. In R. L. Linn (Ed.), *Educational measurement* (pp. 335-366). New York: Macmillan.

Natriello, G., McDill, E.L. & Pallas, A.M. (1990). *Schooling disadvantaged children: Racing against catastrophe*. New York: Teachers College Press.

New Jersey State Department of Education. (1996). *Core curriculum standards*. Trenton, NJ: Author.

Porter, A. (1999). Creating a system of school process indicators. *Education Evaluation and Policy Analysis*, *13*(1), 13-29.

Quality Counts 2000. (2000). *Education Week*, *XIX*(18).

Raymond Arthur Abbott, et al. v. Fred G. Burke, Commissioner of Education, et al., 643 A.2d 575, 580 (1994) (Abbott III)

Raymond Arthur Abbott, et. al. v. Fred G. Burke, Commissioner of Education, et. al., 693A.2d 417, 428 (1997) (Abbott IV).

Rothman, R. (1995). *Measuring up: standards, assessment, and school reform*. San

Francisco, CA: Jossey-Bass.

Schmidt, W. H., McNight, C. C., Raizen, S. A. (1997) *A splintered vision: an investigation of U.S. science & mathematics education*. New York, NY: Kluwer Academic Publishers

Smith, M. & O'Day, J. (1991). Systemic school reform. In S. Fuhrman & B. Malen (Eds.) *The politics of curriculum and testing.* (pp. 233-67). Bristol, PA: Falmer Press.

Smith, M.L. (1991). Put to the test: The effects of external testing on students. Educational Researcher, 20(5), 8-12.

Smith, M.S., Fuhrman, S.H. & O'Day, J (1994) National curriculum standards: Are they desirable and feasible? In Elmore, R.F. & Fuhrman, S.H. (Eds.) *The governance of curriculum* (pp. 12-29). Alexandria, VA; ASCD.

Wiley, D.E. & Yoon, B. (1995). Teacher reports on opportunity to learn: Analyses of the 1993 California Learning Assessment System (CLAS). *Educational Evaluation and Policy Analysis*, *17*(3), 355-70.

About the Authors

William A. Firestone
Center for Educational Policy Analysis
Rutgers University

Email: wilfires@rci.rutgers.edu

William A. Firestone is Professor of Educational Policy; Chair of the Department of Educational Theory, Policy, and Administration and Director of the Center for Educational Policy Analysis. His research on the effects of both testing and professional development on teachers has appeared in the American Educational Research Journal, Educational Evaluation and Policy Analysis, and Kappan. His most recent book is *From Cashbox to Classroom: School Finance Reform and Educational Change in New Jersey* (with Margaret E. Goertz and Gary Natriello).

Gregory Camilli
Rutgers University

Email: camilli@rci.rutgers.edu

Gregory Camilli is Professor, Department of Educational Psychology, at the Rutgers Graduate School of Education. His areas of research interest include psychometric issues in educational policy, meta- analysis, and differential item functioning. Dr. Camilli is a member of the editorial Boards of *Educational Measurement: Issues and Practice, Educational Policy Analysis Archives,* and *Educational Review*. He is a regular reviewer for *Applied Measurement in Education, Journal of Educational Measurement, Psychometrika,* and *Psychological Methods*, among others. As a member of the Technical Advisory Committee of the New Jersey Basic Skills Assessment Council, he provides expertise on testing and measurement issues to the state's assessment program.

Michelle Yurecko
Center for Educational Policy Analysis
Rutgers University

Michelle Yurecko, a Ph.D. candidate in Educational Psychology and a research associate at the Center for Educational Policy Analysis. She is a statistician and focuses on educational testing and measurement.

Lora Monfils
Center for Educational Policy Analysis
Rutgers University

Lora Monfils, a doctoral candidate in Educational Statistics and Measurement, is a research associate at the Center for Educational Policy Analysis and a mathematics educator. Her research interests concern large-scale assessment and modeling differential educational outcomes.

David Mayrowetz
Center for Educational Policy Analysis
Rutgers University

David Mayrowetz is a doctoral candidate in the Department of Educational Theory, Policy and Administration, Rutgers University. His interests include policy implementation, inclusion of students with disabilities, and assessment reform. He is the co-author, with William Firestone of "Rethinking "High Stakes:" Lessons from the US and England and Wales" (Teachers College Record, forthcoming) and with Carol Weinstein, of "Sources of Leadership for Inclusive Education: Creating Schools for All Children" (*Educational Administration Quarterly,* September 1999). He will be joining the Policy Studies faculty of the University of Illinois at Chicago in January 2001.

Appendix A
New Jersey's Core Curriculum Content Standards

Mathematics:

1. All students will develop the ability to pose and solve mathematical problems in mathematics, other disciplines, and every day experiences.
2. All students will communicate mathematically through written, oral, symbolic, and visual forms of expression.
3. All students will connect mathematics to other learning by understanding the interrelationships of mathematical ideas and the roles that mathematics and mathematical modeling play in other disciplines and in life.
4. All students will develop reasoning ability and will become self-reliant, independent mathematical thinkers.
5. All students will regularly and routinely use calculators, computers, manipulatives, and other mathematical tools to enhance mathematical thinking,

understanding and power.

6. All students will develop number sense and an ability to represent numbers in a variety of forms and use numbers in diverse situations.
7. All students will develop spatial sense and an ability to represent geometric properties and relationships to solve problems in mathematics and in everyday life.
8. All students will understand, select, and apply various methods of performing numerical operations.
9. All students will develop an understanding of and will use measurement to describe and analyze phenomena.
10. All students will use a variety of estimation strategies and recognize situations in which estimation is appropriate.
11. All students will develop an understanding of patterns, relationships, and functions and will use them to represent and explain real-world phenomena.
12. All students will develop an understanding of statistics and probability and will use them to describe sets of data, model situations, and support appropriate inferences and arguments.
13. All students will develop an understanding of algebraic concepts and processes and will use them to represent and analyze relationships among variable quantities and to solve problems.
14. All students will apply the concepts and methods of discrete mathematics to model and explore a variety of practical situations.
15. All students will develop an understanding of the conceptual building blocks of calculus and will use them to model and analyze natural phenomena.
16. All students will demonstrate high levels of mathematical thought through experiences which extend beyond traditional computation, algebra, and geometry.

Science:

1. All students will learn to identify systems of interacting components and understand how their interactions combine to produce the overall behavior of the system.
2. All students will develop problem-solving, decision- making and inquiry skills, reflected by formulating usable questions and hypotheses, planning experiments, conducting systematic observations, interpreting and analyzing data, drawing conclusions, and communicating results.
3. All students will develop an understanding of how people of various cultures have contributed to the advancement of science and technology, and how major discoveries and events have advanced science and technology.
4. All students will develop an understanding of technology as an application of scientific principles.
5. All students will integrate mathematics as a tool for problem-solving in science, and as a means of expressing and/or modeling scientific theories.
6. All students will gain an understanding of the structure, characteristics, and basic needs of organisms.
7. All students will investigate the diversity of life.
8. All students will gain an understanding of the structure and behavior of matter.
9. All students will gain an understanding of natural laws as they apply to motion, forces, and energy transformations.

10. All students will gain an understanding of the structure, dynamics, and geophysical systems of the earth.
11. All students will gain an understanding of the origin, evolution, and structure of the universe.
12. All students will develop an understanding of the environment as a system of interdependent components affected by human activity and natural phenomena.

Appendix B
Content Area Topics From The Teacher Survey

Mathematics:

1. Paper and pencil mathematical operations with whole numbers (adding, subtracting, multiplying & dividing)
2. Doing mental math operations with whole numbers (adding, subtracting, multiplying & dividing)
3. Estimation (magnitude, results of computation, measurement)
4. Place value relationships (whole numbers, decimals)
5. Adding and subtracting decimals via paper and pencil
6. Identification of geometric figures
7. Area and Perimeter
8. Fraction Concepts (Fractions as parts of a whole, equivalency)
9. Operations with Fractions (addition, subtraction)
10. Measurement (customary, metric)
11. Probability
12. "Dealing with data" (collecting, organizing, analyzing and displaying data)
13. Statistics
14. Graphing
15. Patterns, functions
16. Open sentences, use of variables
17. "Discrete math" (Combinations, puzzles, optimization, classification, algorithms, networks, tree diagrams)

Science:

1. Understanding natural and man-made systems (recognizing systems, identifying parts)
2. Investigative skills (observing, classifying, dealing with data)
3. Using mathematics (measurement, estimating, counting)
4. Nature and history of science & scientists
5. Selecting and using tools
6. Needs of living things/Life systems
7. Habitats, ecosystems, & adaptation
8. Features and classifications of plants and animals
9. Structure and physical properties of matter
10. States of Matter: Solid, liquid, gas (heating and cooling)
11. Forces, motion & energy

12. Invisible forces (gravity, electricity & magnetism)
13. Earth Materials: Rocks, soil, fossils
14. Weather and climate
15. Earth, moon, sun system
16. Stars and galaxies
17. Humans and the environment

Appendix C
Sample ESPA Items

Traditional Mathematics Item:

Find the exact answer: 110 + 70

1. 18
2. 81
3. 180
4. 810

Newer Mathematics Item:

Mr. Jones gave each of the students in his class a one-ounce box of raisins. When the students opened the boxes and counted the raisins, they found different amounts. The tally sheet below shows their results.

Number of Raisins	**Tally**	**Frequency**
10	I	1
11	II	2
12	III	3
13	IIIII	5
14	III	3
15	II	2

Construct a bar graph to represent the students' findings on the grid in your answer booklet. Be sure to label your graph completely.

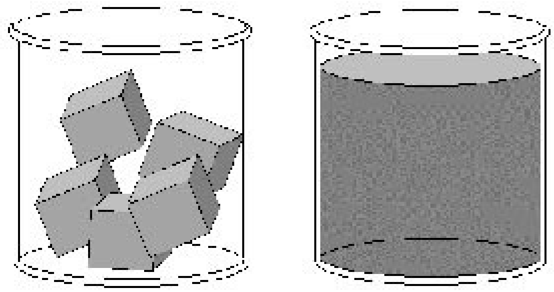
Traditional Science Item:

Which thing does a living duck do that a toy duck does not do?

1. Floats on water
2. Breathes air
3. Makes a sound
4. Sits still

Newer Science Item:

Victor has two glasses. One glass is filled with ice cubes and the other is filled with water. Give three ways the ice and water are different.



The World Wide Web address for the *Education Policy Analysis Archives* is **epaa.asu.edu**

General questions about appropriateness of topics or particular articles may be addressed to the Editor, Gene V Glass, glass@asu.edu or reach him at College of Education, Arizona State University, Tempe, AZ 85287-0211. (602-965-9644). The Commentary Editor is Casey D. Cobb: casey.cobb@unh.edu .

EPAA Editorial Board
