
Public Hearing

before

SENATE EDUCATION COMMITTEE

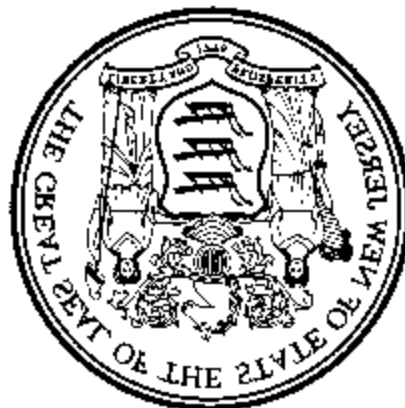
“Testimony concerning the Elementary School and Grade Eight Proficiency Assessments”

LOCATION: Committee Room 6
State House Annex
Trenton, New Jersey

DATE: May 31, 2001
1:00 p.m.

MEMBERS OF COMMITTEE PRESENT:

Senator Robert J. Martin, Chairman
Senator Joseph A. Palaia, Vice-Chairman
Senator William L. Gormley
Senator Byron M. Baer
Senator Shirley K. Turner



ALSO PRESENT:

Darby Cannon III
Office of Legislative Services
Committee Aide

Christine Shipley
Senate Majority
Committee Aide

George LeBlanc
Doug Wheeler
Senate Democratic
Committee Aides

Hearing Recorded and Transcribed by
The Office of Legislative Services, Public Information Office,
Hearing Unit, State House Annex, PO 068, Trenton, New Jersey

TABLE OF CONTENTS

	<u>Page</u>
Vito A. Gagliardi Sr., Ph.D. Commissioner New Jersey Department of Education	2
J. Edward Doolan Director Standards and Professional Development Division of Academic and Career Standards New Jersey Department of Education	33
William A. Firestone, Ph.D. Professor of Educational Policy, and Director Center for Educational Policy Analysis Rutgers University	35
Lora Monfils Senior Research Analyst Center for Educational Policy Analysis Rutgers University	55
Mary O'Malley Vice President Local Initiatives Prudential Insurance Company	67
Dana Egreczky Vice President Workforce Development New Jersey Chamber of Commerce	73
Douglas Groff Chairman Statewide Assessment Advisory Panel New Jersey Association of School Administrators	83
Robert L. Bumpus Chairman	

TABLE OF CONTENTS (continued)

	<u>Page</u>
Curriculum and Instruction Committee New Jersey Association of School Administrators	83
Trudy Doyle Co-Chairperson Ad-Hoc Committee on Assessment of Core Curriculum Content Standards New Jersey Association of School Administrators	84
Len Mitnaul Teacher Woodrow Wilson Elementary School	98
APPENDIX:	
Presentation submitted by William A. Firestone, Ph.D.	1x
Testimony and position paper submitted by Dana Egreczky	27x
Statement submitted by Garden State Coalition of Schools	43x
Testimony submitted by Christine Carney Assistant Director/Lobbyist Governmental Relations New Jersey School Boards Association	44x
Testimony submitted by Linda M. Kassekert	

Associate Director
Government Relations
New Jersey Education Association

46x

TABLE OF CONTENTS (continued):

APPENDIX (continued):

Page

Testimony
submitted by
New Jersey Principals and Supervisors Association

50x

Report
submitted by
Peter J. Calvo and
Louise Murray-Hoffman
Co-Chairpersons
Ad-Hoc Committee on Assessment of
Core Curriculum Content Standards
New Jersey School Boards Association

52x

rs:1-106

SENATOR ROBERT J. MARTIN (Chairman): This is a hearing of the Senate Education Committee.

We were waiting patiently for Senator Turner. She was the one who first asked the Committee to have a special session devoted to taking a look-see at where we are with the testing -- mandatory testing with the State of the fourth-, eighth-, and eleventh-grade levels, but especially at the fourth-grade level, where the new testing has been in place now for a couple of years. And I dare say no one would be surprised if I suggested there have been at least some concerns raised about the adequacy of the testing, the length of the testing, other problems associated with it.

I can speak personally and say that we have heard, in our legislative office, and I think this is true with many of my colleagues, complaints that have come in from teachers who have indicated that it is very time consuming, difficult, and there has been some issues along those lines. Whether that's just part of the difficulty of change, or whether it's more systemic than that, we thought it was appropriate to have a hearing simply devoted to the issue.

We could have had this, perhaps, at all levels of the testing, but the most concerns we've heard deal with the fourth-grade testing. So, although we've more or less limited this hearing to that topic, those more generic areas such as eighth-grade -- the eighth-grade test and the eleventh-grade test are also -- if they're part and parcel of the general theme, we would also welcome sort of overarching comments related thereto.

With that said-- And we acknowledge Senator Turner's role in this. She had been the one who had first asked that we spend some time and devote some attention to this theme.

I would like to call up the Commissioner of Education, Dr. Vito Gagliardi, who does have some time constraints, but he was kind enough, some months ago, when he first became Commissioner, to agree to talk specifically about this issue. And seeing him here personally to talk about that is refreshing. And we look forward to his commentary.

Commissioner, the floor is yours.

COMMISSIONER VITO A. GAGLIARDI SR., Ph.D.:

I was just reading that I had to push the black button, Senator. (referring to PA microphone) And I did indeed do that.

Thank you.

SENATOR MARTIN: That's the first step in the process.

COMMISSIONER GAGLIARDI: That's part of the test, by the way. We will include the black button at some other point, but reading the directions is always a good idea.

SENATOR MARTIN: From what I understand, though, our instructions are shorter and simpler than those posed on fourth-grade students. (laughter)

COMMISSIONER GAGLIARDI: That's one for you, Senator. (laughter)

Good afternoon, Mr. Chairman and members of the Senate Education Committee. I have come back at your request to discuss the Elementary School Proficiency Assessment and the Grade Eight Proficiency Assessment, which I will refer to as the ESPA and the GEPA. One of these days, we're going to run out of alphabet soup, and then we'll have to always say the

things we mean as opposed to the acronyms. But for today, I'll use ESPA and GEPA.

As most of you know, I've spent 37 years in education, from classroom teacher to now being Commissioner, and I've had more than firsthand experience in all of New Jersey's testing programs, and, I believe, the effect that it has on classroom instruction. I guess I've experienced all sides of the issues.

Since New Jersey has rigorous academic standards that students are expected to meet, there must be tests aligned with our standards to show progress toward meeting them.

Some districts ask why their commercial testing programs are not as good as the State assessments. For the purpose of determining progress of students on achieving New Jersey's standards, it is essential to have tests that are directly aligned with our standards. Commercial tests are not aligned with our standards, nor are national assessments. There may be some commercial tests that are becoming more rigorous, but they are not yet up to our standards.

The State's assessment program for Grades 4, 8, and 11 is an essential tool for measuring to what extent students have learned the concepts in the core curriculum content standards. There was much anxiety in the field two years ago over the new ESPA, so the Department conducted an extensive outreach effort to allay concerns of parents and educators.

On our Web site, the Department of Education has made available sample forms of test questions for the GEPA in language arts literacy and math, and the ESPA in language arts literacy, math, and science. To explain the testing programs, and what students and parents should expect, the Department

developed parents' guides in both English and Spanish for the ESPA and GEPA. Both were distributed in print, and they also are available on the Web site.

In addition, on the NJPEP page of the Department's site, our professional development portal for teachers, there is a tutorial on how to interpret test results. And I would like to add that I have experienced that particular Web site myself, and I endorse it, and I think it can be very helpful.

In response to the concerns that were expressed initially, the tests have been modified. Test items have been reviewed, and the overall length of the ESPA and GEPA have been shortened. We think we have arrived at a point where the tests meet the needs of both the State and local districts to determine how students are doing in acquiring the concepts in the standards. As Commissioner, I have pledged to all stakeholders that I will continually evaluate all of our operations and make change where it is needed.

Now, because of the concerns about the quality and length of the tests expressed by the public and the business community, especially through the Business Coalition of Education Excellence of the New Jersey Chamber of Commerce, the Department of Education asked the nationally acclaimed organization, Achieve, Inc., to evaluate both the ESPA and GEPA. Achieve, as most of you know, is an independent, bipartisan, nonprofit organization formed in 1996 by governors and corporate CEOs who share a belief that high academic standards and accountability for performance can push our schools and students to much higher achievement.

I may add that this is something new, at least to me. In the past, when I was a school administrator, all we heard were complaints of the business community in terms of how poorly our students were doing. This is a new era,

especially in New Jersey, where the business community is offering a helping hand to identify what is good and what needs to be improved. And I have shared with all of those individuals in the business community my appreciation for their continued involvement.

Last year, the Department requested that our assessments be evaluated by Achieve's benchmarking process, by which our tests were measured against our State standards to determine if we are truly assessing what we expect our students to know and be able to do. Such an evaluation gives the State objective, credible, concrete recommendations for ways to improve our standards and assessments.

Both the ESPA and the GEPA have undergone rigorous scrutiny by Achieve. The assessments in language arts literacy and math were found to be superior in quality among those reviewed. Items in both content areas were deemed to be challenging, and reading passages were termed rich. The level of challenge was found to be generally high, and the use of open-ended items was seen as good, because, in their words, they elicit knowledge and skills not always tapped by multiple-choice questions.

The Department is currently submitting its science standards, curriculum frameworks, and assessments for benchmarking. Results are expected by the end of the summer.

In the first administration of the ESPA, there were too many variations among the scores on the writing tasks in language arts literacy. This discrepancy triggered a review of these items. Concern was voiced about the length of the passages and the application of the standards used to judge writing samples.

After the year 2000 administration of the ESPA, the Department took several steps intended to address public concerns about the test. Last year, the Department recalculated the language arts literacy scores for 1999-2000, because the scores were not consistent with the result found in the math and science exams. This action was taken so that the scores would more accurately reflect the performance and abilities of fourth-graders.

In addition, the Department is conducting an intensive review of the ESPA language arts literacy scoring system to ensure it is in sync with curriculum expectations, and we are making more item-specific information available to districts.

For example, we have made available to districts an actual writing task from the 2000 ESPA, and each student's written response with sample score points, so teachers and administrators can see how their students fared and what could have been done to improve scores.

The Department has also taken steps to reduce the length of the passages, as well as the overall assessment. The overall length of the ESPA language arts literacy test has been reduced to approximately three hours over a two-day period. The total time required for all sections of the ESPA, as currently organized, is six hours in Grade 4 over a week's time, and one and a half hours in one day in Grade 5.

The Department is cognizant of the need for children to stretch, move, and refocus in order to perform optimally. With one exception, a student is not expected to devote longer than 30 minutes to any one task. The one exception is that students are given 50 minutes for the writing tasks.

During the first administration of the ESPA, test anxiety was high among teachers, administrators, and parents. This is always true of new assessments in this state or any other state. Following the administration of the ESPA this spring, the anxiety seems to have abated significantly, as it will with each test year.

It is difficult to get direct feedback from the students, but a recent article published in *The Star-Ledger* reports on interviews with fourth-grade students who had recently taken the ESPA tests. By and large, the students did not reflect the level of stress that many people had attributed to them. The students found the multiple-choice questions accessible. The open-ended questions were described as making you think a little more, but the students felt these questions were also within their reach. The writing task that allowed them to create their own monster was engaging and enjoyable to the students interviewed. In general, the 2001 ESPA was not seen as a negative experience for the students.

Another issue that always arises with a new assessment program is the charge that teachers will teach to the testimony. That criticism was expressed when we went to school and even before that.

William A. Firestone, Professor of Educational Policy at Rutgers and Director of the Rutgers Center for Educational Policy Analysis, recently presented findings of a study on the impact of New Jersey testing on teaching in a series of papers to the American Educational Research Association. I met Professor Firestone today for the first time. And obviously, I've read portions of his papers.

The study found that the assessments had “encouraged teachers to try out more inquiry-oriented instruction. Such instruction included placing a greater emphasis on problem-solving, having students explain their thought process, assigning more writing and making greater use of hands-on materials.” The research suggests that “the tests appear to be sensitizing teachers to new instructional practices, even if it is not leading to major change.”

Although the study implies that so-called teaching to the test does not productively use classroom time to the fullest, we do not fully agree. It has always been our belief that if the standards are strong enough and our assessments are of high enough quality to accurately measure student progress toward teaching the standards, then teaching to those tests will enable students to attain the standards that have been set.

When ESPA and GEPA were aligned with the State’s core curriculum content standards, they were constructed to require students to think through a multistep process to answer the questions. The Rutgers research contradicted the assumption that teaching to the test means almost always driving instruction to the lowest level. Dr. Firestone found the tests to be encouraging teachers to consider more challenging approaches that incorporate problem solving and critical thinking.

However, the study also found that the tests also “could be encouraging more conventional and drill-oriented instruction in the State’s poor urban districts, where such teaching is already more prevalent.” While the tests are generally encouraging teachers to try new methods, these changes are not substantial enough.

The results of this study suggest to the Department that we are on the right track by making the test require higher level thinking, but it is a clear mandate for us to provide more professional development to help teachers upgrade classroom practices for developing students' thinking skills and reduce drill-oriented instruction. We are working to help districts with this aspect, especially our *Abbott* districts, where the greatest need for change exists. We want all of our students to acquire not only the knowledge required in the standards, but the skills, as well.

We want to ensure the validity and reliability of the tests in the areas currently assessed and to ensure the satisfactory achievement of all New Jersey students in these four areas. Additionally, in accordance with the requirement of the Administrative Code-- Just this past month, on May 16-- Well, it's still this month. I thought today was June 1. But on May 16, I charged a broad-based review committee with the task of evaluating the core curriculum content standards.

The committee's major goal is to review, with a critical eye, the standards in each of the content areas and follow these two steps: to reflect upon the thoughtful recommendations of external reviewers such as Achieve, as well as those educators in New Jersey who have been utilizing the standards for the last five years, and to make whatever revisions are necessary to improve the quality, as well as the clarity, of the standards.

It is not the committee's charge to expand the standards to encompass the totality of learning in a content area. The standards should continue to focus on the essential core of learning in each subject area that is most important for all students to understand and be able to use.

The charge is to continue the clear intent of the original standards, to emphasize a depth of understanding that challenges students to apply and use knowledge rather than emphasizing surface coverage and recall.

Districts are urged to continue to implement curriculum and instruction in the areas that are currently not assessed. Focusing attention and resources on all seven content areas should continue to be the guiding force in education as we face the challenge to provide all students in New Jersey with a thorough and efficient education, one that will prepare them to become productive citizens of the 21st century.

New Jersey understands that continuous improvement of standards and tests is necessary to keep pace with changing requirements in society. The revision of the standards and subsequent review of the assessments provide a unique opportunity to reflect on those aspects that have really served our students well and adjust those that could be strengthened.

The Department staff also has been sensitive to concerns about how the test results are reported. District administrators have expressed a desire for assessments that are capable of generating meaningful information that guides districts in their continuous efforts to improve student achievement, as well as a shorter time frame between the administration of the test and the reporting of results.

The Department is exploring several possible alternatives in the effort to address these issues. We are mindful of the need to assure the validity and reliability of the assessments. When the number of test items evaluating a particular cluster of skills is small, the data for this area is not reliable for the purposes of doing an analysis of the school's strengths and weaknesses. If we

include sufficient items in each cluster to allow the reporting of results that are valid and reliable for district analysis, it would require additional time. We are trying to establish a balance and find a solution that will meet everyone's needs.

We are working with our contractor to initiate procedures that will allow results to be returned to the districts in the minimum time feasible while maintaining reliability in the scoring of the open-ended items. To sacrifice quality in the effort to save time would compromise our high standards, so we must proceed carefully.

It is our expectation that the GEPA results will be sent out to the school districts during the week of June 11. The math and science results for the ESPA will be forwarded to the districts the week of August 13. And the ESPA language arts literacy results will not be ready for distribution to districts before mid-September. This is because we will be setting new cut scores for partially proficient, proficient, and advanced proficient performances on the exam. We are taking this step because of experiences last year which led to a rescoring of ESPA's language arts literacy component.

The testing program also may be impacted by the proposed legislation before Congress that mandates testing in language arts and math in Grades 3 through 8. When the debate is over and the education agenda is finalized, we will study the requirements carefully, consider the impact on our statewide program, and assess viable strategies and their implications at that time. Whatever the outcome, the message is clear: That we must be able to show accountability in the form of student achievement for the investments we make, whether from the Federal government, the State, or the local community.

Over the past seven years, the Legislature has made an extraordinary commitment to the public school children of New Jersey. State aid to education has gone from \$4.4 billion in Fiscal Year '95 to \$7.4 billion for the 2001-2002 school year. The State's 30 *Abbott* districts will receive a total of \$2.9 billion, the most they have ever received, and 49 percent of the entire amount of State aid to be distributed. This includes the more than \$429 million in *Abbott v. Burke* parity remedy aid that *Abbott* districts will receive next year in order to continue parity in per-pupil spending at a level equal to the wealthiest districts in the state.

With that investment, the public expects accountability for results. We have pledged to educate all children, and the major task that the State has for determining how successful we have been is the statewide assessment program.

Many view tests as intrusions on the learning process, but tests are an integral part of teaching every single day. If teachers do not monitor student progress in acquiring concepts by various types of formal and informal assessment methods, they have no way to know what each student has mastered.

On the State level, we must give the districts feedback on how well all of their students are progressing in acquiring the concepts and skills in the core curriculum content standards required by the State. It is through skillful use of assessments that we can assure that all students are advancing appropriately. We added the ESPA to the GEPA as periodic diagnostic reports to districts that show whether students are making sufficient progress to assure

that they will know and be able to do what is measured in the HSPA, New Jersey's high-stakes test.

The Department of Education fully acknowledges that a successful assessment program evolves through the evaluation of good feedback from many sources and making the appropriate adjustments. We believe we have followed that process, and we will continue to do so.

I will be happy to take your questions, Mr. Chairman.

SENATOR MARTIN: Thank you.

Just a couple of questions, and then the Committee will ask some.

And just so that folks know, Dr. Firestone was alluded to by Commissioner Gagliardi. He is here, and he and his associate have prepared a presentation, which is why we have the screens. So he will go into detail about their review of New Jersey's testing. So, if the question is related to that, I think we should hold for Dr. Firestone.

Commissioner, you indicated that the amount of time for the fourth-grade test has been reduced significantly. Are we-- Just clarify something for me. Are we going to have an additional test, though, as other segments of the core curriculum standards come in, or is that not--

COMMISSIONER GAGLIARDI: I'm wrestling with that right now. It would seem logical for the students, the State, for all of us, that if we're looking at the standards, and it's going to take about a year, we would have to take another look at the assessments.

My major concern, Senator Martin, is that I don't want the Department of Education to make the mistakes -- they were smart mistakes, by the way -- that were made in the past. So I'm wrestling right now, as we speak,

about how we're going to be dealing with those other items that require an assessment.

SENATOR MARTIN: My other question is--

COMMISSIONER GAGLIARDI: I guess my answer is maybe.
(laughter)

SENATOR MARTIN: And I think at this time maybe that's the best answer, because I don't think -- speaking intuitively, I'm not sure the State should commit to having some testing, especially at the same time as these, being administered as part of that bulk of mass that is done. If it could be done in some other way and through portfolios and other kinds of things, maybe there might be some other avenues to be explored. I think we have to have some kind of comprehensive measurement at some point, but I'm not sure it has to take this sort of traditional, this-is-the-week-of-the-testing kind of process.

That aside, one of the issues that seems to be fundamental, and you alluded to it, is teacher preparation. And there have been some steps taken by the Department, but I've also heard from others -- some professionals in the field, who have suggested that we really need to do more for teachers to be able to have them not teach to the test, but really prepare them to handle the core curriculum content standards. And that's been especially stressed in some of the *Abbott* districts, particularly the -- some of the -- especially the three State takeover districts, where we know that, for example, Newark has been advertising as having a shortage of teachers -- having a difficulty in finding teachers. I think they're recruiting in India and through the military and through other locations, which I'm not passing any judgment on. It just seems to emphasize a point that it is difficult to be able to have teachers, let alone are the

teachers then going to be prepared to engage in this more sophisticated process of being able to provide for this kind of instruction.

So my question is, what is the Department doing to try to help teachers deal specifically with the core curriculum content standards?

COMMISSIONER GAGLIARDI: I think we are attempting to meet that particular need, and have been in the two months -- it seems like two years -- that I've been in office here. But I think we're talking about a number of things. We're talking about the preservice, and that is people who think they want to be teachers. So we're certainly talking with the higher ed people.

In addition to that, obviously, the in-service, the professional development -- whereas I have been referring to it in some of the speeches that I have been making -- thinking of it as developing professionally, as other professions do-- So we are very interested in doing that. And certainly, Governor DiFrancesco is very concerned about that, too.

So we're addressing the manner in which teacher candidates -- those students that think they want to be teachers -- the manner in which they're being instructed while they go to college. And we're certainly taking a look more closely at the professional development of teachers that are presently on staff so that they feel better about this.

I mentioned, in my prepared comments, NJPEP. It's a virtual academy that, I felt, if I was going to speak about it, I ought to use it. And I'm amazed that-- And I suggest to all of you, if you have about 20 minutes to get a feel for it, get on our Web site and just click onto that. You'll be amazed at the professional development that is available on that particular site. And the division that is responsible for the testing and all the rest are doing a great job

in terms of making it available to not only teachers that are presently on staff, but those that want to be teachers, and I think it has to be handled--

Now, unfortunately for the teacher candidate, it's going to take some time before they come through the pipeline, but if we don't start soon, we'll never begin. So I think it's important. And I think the cooperation that we're receiving from the colleges and the universities right here in New Jersey is just great.

SENATOR MARTIN: Senator Palaia.

SENATOR PALAIA: Thank you, Mr. Chairman.

Commissioner, it was a very interesting testimony -- a couple of things.

I went to the highest source I could find about this year's fourth-grade testing, the very highest I could go to.

SENATOR MARTIN: Your grandson.

SENATOR PALAIA: My daughter. (laughter) I went right to the top. And she told me that the testing this year was much better. And again, she had been critical, obviously, of all the nuances with the length of time and the marking and all this, but she said that this year, it was much better than it had been in the past. So that's from the trenches -- people who are there.

The comment you make about teaching to the test-- As you said, I've been in education 33 years, too. Let me tell you, there's nothing wrong with that whatsoever, nothing wrong with that. They're getting the same feel for the curriculum that has been proposed. And I think those who say it's wrong-- I don't agree with that at all.

One last question. How do we deal with special ed students, Commissioner, as it relates to this test? I know that was a problem from years ago. Do they take it? Don't they take it? And if they do, is it marked against you on the State's average? How do you deal with that, besides very carefully?

COMMISSIONER GAGLIARDI: Yeah, very carefully. I think we're still reviewing that, Senator. And let me tell you why that's important. And I'm sure this isn't done intentionally. But one could question whether or not a youngster who was identified as having learning problems with one of the labels is -- apt to do -- would eliminate that youngster from the scores. And I think we have to be very careful that we don't play into that particular hand.

I think, right now, I know that both divisions that are responsible for that are looking at that. I don't know specifically if we made a decision on that yet, but I know it's a major concern for many reasons, because I think there are youngsters, regardless of their particular classification, who could do well on that test.

And maybe I might defer to one of my staff, who could give you a more definitive answer. But it is a concern that we've discussed over and over. And it's a very good question. I wish I had the complete answer for you right now. But we'll get it for you.

SENATOR PALAIA: Senator Gormley wanted to follow up.

SENATOR GORMLEY: We did a tour -- a special services school district in Atlantic County, and we've also done other school tours, and everybody does that on the Committee.

SENATOR MARTIN: Are you on your-- (referring to PA microphone)

SENATOR GORMLEY: I'm sorry.

I was mentioning we did tours of the special services school district in Atlantic County, and we've had a lot of correspondence back and forth on this. What happened at the special services school district -- the direction they got-- And this is one of those games -- it's Federal, it's State, and the local district is really put in an incredibly difficult situation.

They had to prepare an exam for the students in lieu of the standardized test. They had to assign the school psychologist -- "Go ahead, write an exam." So they're put in the predicament of taking a valued professional -- taking an enormous amount of time from that valued professional, because there is an acknowledgment that unless you take the test, an alternate test can be provided. It is somewhat cumbersome, obviously. It is time consuming of someone who's very valued in a special services school district, and it's something where a professional, when assigned something, really struggles with it.

So is there any way we can do something about this, because what they get hit with is -- now, that's a Federal mandate, but it's a situation that is really a-- You're taking away productive time -- giving them an assignment to create a test that really isn't of much value given certain levels of disability. I don't mean that in a condescending fashion. I mean, that psychologist's time is better spent interacting with that child. And it was-- You've heard this from other people, I assume.

COMMISSIONER GAGLIARDI: Oh, absolutely.

SENATOR GORMLEY: Now that's something we've got to do something about.

COMMISSIONER GAGLIARDI: Yes. And let me add that-- And you probably know that it's not just a K-12 situation. You probably know neighbors or friends or maybe even those in your own family where they graduate high school and they go to a particular university or college. Allowances are made now, thank goodness, for those youngsters that are identified by the college as having some learning disability. And there are some allowances made, for instance, for tests that are given at--

So I guess the point I want to make is it's not just a pre-K -- 12 problem. And I'm not one to just point our finger at Washington and say, "The feds made us do it." That's not fair. If it's a good thing, obviously, we should be looking more closely at it. And I'm just telling you that we are. I don't know if I had made, with our staff, a complete decision on it.

SENATOR GORMLEY: So our predicament is that it's ongoing in the special services districts.

COMMISSIONER GAGLIARDI: Yeah.

SENATOR GORMLEY: And it creates a hole in the budget because of the hours of those professionals. And you're aware of this.

But if it is a mandate, could there be a team of people from the State who do it or provide -- because the examination sounds like window dressing to me to say that they did an examination. And it's just an untenable situation, but it's taking up a lot of time. And we did send--

We've been back and forth about this -- not you personally.

COMMISSIONER GAGLIARDI: No, no, I know.

SENATOR GORMLEY: But it really should be addressed now, because it's real, and it affects the districts.

COMMISSIONER GAGLIARDI: Just let me respond in this way. I know one of the concerns. And it was meant to be a good thing. You're all familiar with the SRA process as a result of a youngster not being able to pass the easier of the high school proficiency assessment.

SENATOR GORMLEY: Yes.

COMMISSIONER GAGLIARDI: And the Department was -- there was fault found there in terms of it's a backdoor approach to the diploma. So I've been aware of that particular concern for a long, long time.

Obviously, the fourth- and the eighth-grade tests are leading to the HSPA, and I was mentioning before I came up here that I was thoroughly amazed, looking at Achieve Inc.'s benchmarking to the best-- And they had charts.

I confess, and maybe you people know this, but there are only nine states that demand that a student pass a particular test in order to be eligible for a high school diploma. Now, there are others that are thinking about it. But I was amazed.

So I think many of the things that-- And I'm not defending anything. This all happened before I got here, you understand. I think the things that have been -- fault -- with the Department of Education is because, in many cases, the Department and the State of New Jersey have been pioneers in many of these things. And obviously, we had nothing to fall back on -- to replicate. So others will learn from our mistakes, hopefully. And we're looking at it-- I don't say every day. There are people in the Department that look at it every day. But whenever the subject comes up, we make certain, wherever we

can, that we're considering all angles of this particular thing, because we know how difficult it is.

SENATOR MARTIN: Just dealing specifically with Senator Gormley's issue-- I think he raises a good point. It's my understanding of the IDEA -- they do -- you can compel students to take the normal testing, or you can create the alternative testing, which, in many cases, is more appropriate to their particular handicap. I think he's on the right track, however, in saying that it really becomes an incredible burden for each district. Maybe we will look to the State.

It's more complicated because every child is entitled to their own individual assessments, so presumably each test would have to be sort of structured -- not saying they couldn't use somebody else's test, but somebody would, at least, have spent the time to make sure that this test was appropriate for that student and their individualized handicaps. So it's an enormous task. It tries to meet some goals while creating problems that, I think, were unanticipated and unforeseen. But I think there could be a way in which resources from the State could assist, at least to a certain extent, the local school districts to-- Otherwise, it just becomes-- It's already a difficult burden trying -- because they're bearing a brunt.

We're trying to move a special ed bill. The funding issue was held in the Appropriations Committee today. I'm not pleased to indicate, although I have some commitment from the Governor's Office that they may -- and the Senate President -- slash -- that they may permit us to move ahead, even though the funding issues are growing in the State of New Jersey -- that we would get some assistance on that. But the special ed areas -- problems seem to cut

through so many issues, and this is just one more example where we're trying to deal with fourth-grade testing and eighth grade and eleventh grade, and yet special ed complicates those matters.

COMMISSIONER GAGLIARDI: Just let me finish on special ed, if I might, Senator.

There probably isn't any one particular item more than special education -- and we generalize -- probably nothing more important than dealing with those particular problems in the State of New Jersey, whether we're talking about curriculum content standards or anything else. And we can never come up with a prototype for testing or anything else. If we understand that early identification and intervention is important, and you've all heard me talk about my concern concerning autism, but for the family and the student, there is no time or money that is enough to deal with those issues. And obviously, there has to be a line.

But when things go, I guess, negative, people would like to point to the State, and that's all of us. But on one hand, I believe that nobody knows those youngsters, or those groups of youngsters better than the local educators dealing with them. And whether we're dealing with IDEA and we talk about funding, and so on and so forth, it's a very sensitive issue for all of us, and I think it requires a great deal of time, and that issue deserves all the time that we can spend on it.

I would like to end my comment on that.

SENATOR MARTIN: Senator Turner and then Senator Baer.

SENATOR TURNER: Thank you, Mr. Chairman. I want to thank you for allowing us to devote this time to this issue.

And thank you, Commissioner, for being here with us.

I became very, very concerned. In fact, I was very troubled by the recent results of the assessments -- the tests that were published in the local newspaper -- to see the glaring difference between the test results of the City of Trenton and the suburban districts. And I find it extremely troubling that you're telling me we're spending roughly 49 percent of education dollars for the *Abbott* districts, the urban districts, and we have those kinds of results.

I can understand we all want accountability, but I am questioning, too, reliability. Can you tell me why do we have the dramatic difference between the test scores when you're putting that kind of money in the urban districts and the kind of test scores that we're seeing from the suburban districts? How do you account for that glaring difference?

COMMISSIONER GAGLIARDI: Senator, I don't know all the answers to that, but I am more concerned about that issue than anything else. The fact that we are about to desegregate the scores -- not only because it's necessary, but because it's also a Federal mandate -- will make that particular picture, I think, even worse than you can imagine.

However, the unfairness of the question is that if I believe in local control, as I have stated forever, I think those issues are dealt with by the local governance, the local board of education, the local administrators. Those are the people that should be able to speak to not only in Trenton-- I'm not picking on Trenton, I'm talking about any school district, whether it's an *Abbott* or a non-*Abbott* district.

The Department of Education doesn't have enough staff to go into every district to review those particular scores to say you should do it here, you should make this improvement there.

If I believe in local autonomy, and trust me, I do, I think those decisions are made best by the local administrators. They should be asking the question of middle management, the classroom teacher. They should be able to provide the answers that not only you require, but we all require. To expect the Department of Education to do this, I think, is probably a difficult chore. I haven't placed a great deal of money into our budget, specific from the new budget. To do that would be unrealistic.

I agree with you. It's a major concern. I don't have the answers as to why it is occurring. I think the local school districts should know the answer. And if they don't know the answer, I think they should find out the answer.

SENATOR TURNER: And even the charter schools-- The test results there were abominable.

COMMISSIONER GAGLIARDI: Absolutely. One of the things that you know, however -- the difference in the authority of the Commissioner is, if a charter school isn't doing what they said they were going to do in the charter school, the Commissioner of Education has the authority to first put them on notice and then shut them down if they're not doing what-- I do not have the authority to do that in all of the public schools in the State of New Jersey.

SENATOR TURNER: And you mentioned, too, that you don't believe that teachers are teaching to the test.

COMMISSIONER GAGLIARDI: I don't believe I said that. What I said is it's probably happening to some extent. Senator Palaia, I think, agreed with the fact that there's nothing wrong with it.

Let me say that there should be more to the instruction than teaching to the test. I think it happens in more communities than it should because of the stress that they feel. Others, though, have seen it as saying, "Well, at least the students are learning that particular thing."

Perhaps this is not a good example, but when we all prepared for the driving test, we knew specifically what it was we were supposed to do. And the person, whether we paid someone or one of our relatives, taught us what it was we were supposed to do-- If there was nothing else, they taught us how to make the K-turn, how to park, how to use the directionals, and so on and so forth.

It's not that I think that it's a bad thing or a good thing. I think there's more to learning than just that. And that's what I meant by my comments.

SENATOR TURNER: Well, I know teachers have told me that they do start teaching to the test in September and October, and they continue it until the time they take the test in the spring. And their real concern is that the State continues to expand the core curriculum, and there's more and more for them to test -- prepare for. And at the same time, we have not lengthened the school day nor the school year. And there's just all -- you know, so much you can do--

COMMISSIONER GAGLIARDI: Yes.

SENATOR TURNER: --in that short period of time. But they're being required to teach more and more and to be held accountable for more and more of the content.

COMMISSIONER GAGLIARDI: I agree with you.

One of the things I found the Commissioner can't do, because if it's a mandate then we pay-- I know, as a result of a report that is requested, that a number of school districts have expanded their day -- extended the day and the school year. And wherever I can, I endorse that. I think it's necessary. I agree with you.

For instance, a city in Newark has extended their school day by 90 minutes. We cannot have the same number of hours in the day and the same number of days in the school year and expect the youngsters to do what it is that is expected of them. And yet, I do not apologize for the standards that we require, because it's necessary. And I think we have to understand that there will be more.

And as we get into this business more and more, we should raise the standards, because I think in many cases if we don't, we're giving the youngsters in certain communities the wrong idea. And that is, we shouldn't allow any youngster to feel that they are incapable of meeting those standards. It just might take them a little more effort. And I think that's the important issue here.

But I agree with you, Senator. I don't think there would be any legislator who would be willing to sponsor a bill to say now the school day should be X amount of hours and so many school days -- say, "Okay, it's a mandate. You have to come up with more money." But I think local school

districts have that option, and I certainly endorse that. But I couldn't agree with you more in terms of the fact that there just isn't enough time.

SENATOR TURNER: One last question regarding the reliability or validity of these test scores: Do you feel that those test scores are accurate in terms of the scores that children in the inner city had versus those in the suburban districts? Is it really evaluating the students accurately in terms of what they know?

COMMISSIONER GAGLIARDI: Let me respond in this way: If the standards are the same, and the assessments are testing those standards, then the answer to that is yes.

SENATOR TURNER: Because I could not figure out why there was so much difference between the scores that the students had in the urban areas in English, in language arts, math, but then they were so low compared to their counterparts in the suburban districts. But then when you looked at the science scores, they were much, much closer. I would think if they had problems with English and language, they would not be scoring this high on science.

COMMISSIONER GAGLIARDI: Right. And you're right. And that was one of the areas of concern last year. And that's why they were rescored.

I think one of the things that the *Abbott* decision has done that I think we will reap rewards in years to come is to identify the fact that the preschool programs for the three- and four-year-olds, I think, will address the situation, whereby our young students in the inner cities will be better prepared

when they enter kindergarten. It's a preparatory, if you will, for kindergarten. And I think that will help us, in terms of literacy, for sure.

So I think that's one of the ways that we will be able to address that particular concern of yours.

SENATOR TURNER: I agree with you 100 percent. That's why I'm very, very concerned about the fact that so few aren't being able to take advantage of that opportunity. We don't have nearly as many of the students who are eligible enrolled in those programs. And I talked to teachers who were involved in the early childhood education programs, and they tell me that they have submitted their proposals for approval back -- December 1, and they still haven't gotten a response from the Department of Education, as far as approval for their plan.

COMMISSIONER GAGLIARDI: I don't know of any that were submitted in a timely fashion -- complete plans -- that have not been approved as we speak.

Just this morning, I addressed a group -- a conference of early childhood people -- about 350. And I have to share with this Committee what I shared with them.

I had to take a time out and ask Assistant Commissioner Margretta Fairweather, "How many people are on your staff, Margretta?" I met with them once, but I don't remember. There's an Assistant Commissioner and six staff people -- six staff people -- in the Department of Education to review all those plans.

Now, they had done an amazing job in the short period of time that that particular Division has put together. I don't make any excuses for them,

but I think they've done a yeoman's job in accepting the applications, reviewing them, sending them back if they were incomplete, and then approving them and letting the people know just how much money they had.

SENATOR TURNER: I think that's probably a concern, not having enough staff, and I think that has to do with planning. And I remember addressing this -- having your predecessor address this point. He felt that they didn't need any additional staffing. But if--

COMMISSIONER GAGLIARDI: Is that right?

SENATOR TURNER: But if you have that kind of backlog, I think it's important that you hire a sufficient number of people so that we won't have to keep people waiting that long to get approval of their plan, because September is right around the corner.

COMMISSIONER GAGLIARDI: I don't know if there are any plans that were submitted -- complete plans that were submitted that are still waiting for approval, but I will check on that when I go back.

SENATOR TURNER: Well, I spoke to a teacher just yesterday who indicated that she was still waiting from the Department of Education.

COMMISSIONER GAGLIARDI: In Trenton?

SENATOR TURNER: No, it's not in Trenton.

COMMISSIONER GAGLIARDI: Well, maybe later you can tell me what it is.

SENATOR MARTIN: Senator Baer.

SENATOR BAER: I, first of all, want to thank you, also, Mr. Chairman, for holding this hearing, and those who suggested this type of hearing, and you, Commissioner, for coming here.

I think this is-- The issue of these tests are bringing into focus what is probably the most difficult issue that we in government are facing and the most serious responsibility we have of all the things before us. And I think it's going to take all of us working together, lending a hand to each other, instead of finger-pointing.

Now, there are a number of things that I wanted to say, and I hope I can get them all out. It may seem to be a jumble of questions and direct and implied questions and statements. I hope you'll bear with me.

I'm focusing on your reference in your statement to analyzing, through these tests -- or whether they are able to be used in analyzing the schools' strengths and weaknesses to the issues that have been raised here relative to special education and whether the State can play a role relative to the unique challenges there, rather than having each district independently try to invent the wheel, and to the issues relating to the differences between suburban and city performance.

First of all, I wanted to ask, relative to this, whether these tests can be used not only as a tool for student -- identifying student performance and to some degree to see what it says about school performance, but also to see how the test can be used in understanding the problem and challenge created.

What I'm referring to is the potential, through a sampling, of seeing not just as we do without sampling, the performance of one school versus another, but whether we find the performance of youngsters that come from a one-parent family are different from those of a larger group -- whether youngsters that come from families where there's very low educational achievement with the parents -- looking at the relationship between income and

performance, or if we don't have the income information, whatever approximation can be derived from the information the census gives us on average, housing costs in the blocks from where they come, and other things that relate to the youngster's home environment that might give us an indication of the degree of challenge that the school system is faced with, which might be useful in determining the extent of compensatory effort that needs to be provided, since the Supreme Court has spoken to the responsibility of the State in terms of providing that compensatory effort, and whether that -- since it is difficult to know all these answers -- and even yourself don't know all these answers -- whether the State, in trying to see what it can learn from things such as this, which I'm using by way of example, but not, by any means, indicating exhaustively, what the potential is of this as a research tool to understand these problems to see what that can tell the State in terms of what's working and not working and how much assistance these school districts need.

Because although they do have a great responsibility, and I think we all favor the idea of as much local autonomy as possible, there may be a question as to how much of this can be left to each district alone to deal with when some districts are challenged far more than others and may not have resources. And the State has some responsibility, according to the Supreme Court, to provide those resources.

I'm also wondering to what extent this information can be used not only as I have indicated in determining what type of compensatory effort needs to be provided to assist the youngsters and the districts that are heavily impacted, but whether that also has to be taken into account in judging the performance of the school district. Because just by going by test scores alone,

without taking into account what the starting place is, a bunch of youngsters may show very great improvement in one district that starts out with a great challenge and very little improvement with another district where the youngsters are very privileged and have very little challenge. And yet the popular impression is that the district, in one case, that's doing the most would be failing, and vice versa.

So I know I've raised a whole bunch of questions, but I think they go very much to our responsibility as a State. And I'm interested to know to what degree you can speak to that now, and to what degree you feel you would like to be able to speak to that later in something that you might prepare for us when you have more time, given the fact that now, with the magic of modern information technology, our words can be repeated back to you or your staff at any time without even the need of a transcript.

COMMISSIONER GAGLIARDI: In one of my--

SENATOR BAER: I knew I had to come up for air somewhere.

SENATOR MARTIN: Maybe you can try and address it relatively briefly.

I just want to say we have this major presentation. I really want to make sure that we devote our attention to that, so I will--

COMMISSIONER GAGLIARDI: Let me just respond, in all fairness.

SENATOR MARTIN: You can have a--

COMMISSIONER GAGLIARDI: Thank you, Senator.

In one of my responses to Senator Turner, I mentioned desegregating the scores. I think it speaks to much of your comment, in terms

of all those particular items that would impact upon a youngster's ability or disability, if you will, in achieving. And that will take place beginning in September.

J. EDWARD DOOLAN: We'll be doing it for next year.

COMMISSIONER GAGLIARDI: Yeah, we're doing it for next year.

One of the things that you didn't mention, which is also something we're looking at, is the mobility of the youngster. And this is being done in certain parts of the United States and doing very well, for instance, in Texas. It's not to identify the student specifically. I mean, we want to safeguard that particular security and so on. But I'm not talking about moving from district to district, but certainly we are--

But in many of our larger communities, the mobility of a youngster from school causes a great deal of concern. I mean, the last three years of my education career have been as a special assistant, as you probably know. I spent a majority of my time in Asbury Park and Camden, two *Abbott* districts, and very different in terms of size and students.

And mobility is a factor, no question, in addition to the things that you mentioned. So we're aware of it. We're attempting to meet that particular need. And it will provide information so that the State can provide the technical assistance necessary so that the local educators can deal with that in a more direct fashion to all those youngsters that I identified.

So I appreciate your comments, and obviously, we have this electronic device that we can recall most of what it is you're concerned about.

And I'm sure, at some point in time in the near future, Chairman Martin will invite me back, and we'll attempt to address those particular issues, as well.

SENATOR BAER: I hope you will undertake that sampling so that you will be able to see these performance differentials not only based on the school population, but other types of populations that you sample.

COMMISSIONER GAGLIARDI: Yes, sir.

SENATOR BAER: Thank you.

COMMISSIONER GAGLIARDI: Thank you.

SENATOR MARTIN: Okay.

Thank you so much for your testimony.

COMMISSIONER GAGLIARDI: Okay.

SENATOR MARTIN: I think we'll defer to have Dr. Firestone at this point.

COMMISSIONER GAGLIARDI: You can give him all the hard questions. (laughter)

SENATOR MARTIN: I'm sure he could answer Senator Baer's question about-- There's been-- I know there's been loads of research about what indicators are such where -- with respect to low income and other factors, as far as what helps -- what suggests problem areas and other factors. But focusing more directly on your role here with your separate study that we talked about briefly, maybe you could tell the Committee who you engage with in terms of your research, and what your Department and your roles are, respectively, with this research, which is being done with the testing of New Jersey.

WILLIAM A. FIRESTONE, Ph.D.: Thank you, Senator.

And I want to thank the Committee for inviting me to speak to you about this important issue.

SENATOR BAER: Mr. Chairman, can I just interrupt a moment so that we know how much notes we're going to have to take?

SENATOR MARTIN: Yeah.

SENATOR BAER: Are we going to get the hard copy of what's being projected?

DR. FIRESTONE: We can make hard copy of this available, yes.

SENATOR BAER: Thank you.

MR. CANNON (Committee Aide): The whole thing is being transcribed, too.

SENATOR BAER: Oh, okay. Thank you.

DR. FIRESTONE: We can certainly make hard copy available.

SENATOR BAER: Thank you.

DR. FIRESTONE: I am the Director of the Center for Educational Policy Analysis at Rutgers, and I want to introduce my--

SENATOR MARTIN: Just before you do-- Well, I probably should have waited, and I apologize.

Darby had mentioned if you have trouble seeing the screen, people can move where they have to in order to see that. And you can move your chairs accordingly or whatever. So we tried to place these in some general convenient place, but if you want to eat your popcorn in front of the screen, you can do so. (laughter)

DR. FIRESTONE: The graphics are good, but they're not that good.

I want to introduce my colleague Lora Monfils, who is an associate on this research project.

We want to spend some time today kind of setting a context for testing in New Jersey before we talk about our study, and then kind of move on to some policy issues we want to talk about. So we want to try to address two questions. The first is, what kind of testing should the State have? And the second is, what else is needed to improve learning in New Jersey?

And let me bring up some background issues. Right now, I think a number of us are looking over our shoulder at the Bush education plan, which is working its way through Congress, which requires, for those of us who are concerned about time, that states test every child in Grades 3 through 8 in math, language arts, and science. And that will be a condition of getting the Federal support for education -- funding support.

There is a risk. I don't think we know how it's going to work out yet. We will certainly have to test more. There is a risk that we will not be testing better, because we will have to sacrifice certain kinds of things. There are a number of positive aspects to New Jersey's testing program now that we may -- there may be under pressure to compromise on if we have to test this much, and I find that a great concern.

I would say that New Jersey is providing an ordinary education to its children at this time. I don't want to be kind of fear-mongering at this point, but if you look at the data about American students in comparison to other--

SENATOR MARTIN: Does that include Rutgers University, by the way? (laughter)

DR. FIRESTONE: No, I'm speaking K to 12 today. (laughter)

American eighth-graders are about average in a 38-country comparative study of math and science. We haven't done NAEP testing of New Jersey students for quite some time, but when we did, New Jersey was not among the top scoring states, although it wasn't doing badly. And we know, as we've already spoken about, that there's great variation within the State in scoring on our State tests from cities to the suburbs.

At the same time, educational expenditures per child in New Jersey are among the top in the country. We seem to be in a permanent kind of little race on this front with Connecticut about what's the top scoring state -- top spending state per pupil. So we're spending a lot of money, we're providing a reasonable education, but not a great education and not an equitable education. And so there's room for improvement.

Meanwhile, nationally, the bar is being set higher. Ideas about what constitutes an effective education-- And I want to just hit a couple of points from the National Council of Teachers of Mathematics' principles for what effective mathematics education ought to be. The NCTM is probably setting -- kind of leading the pack with regard to content area associations and setting standards for states to use in guidance in developing their own standards.

But the NCTM is saying that students must learn mathematics with understanding, and they should be actively building knowledge from experience. It's not enough to memorize the procedures anymore, it's important to understand mathematical principles. Second, excellence in mathematics education requires equity, going back to Senator Turner's point, high expectations, and strong support for all kids. And third -- I didn't pick out all of their principles, but the ones that are probably most relevant -- assessments

should support the learning of important mathematics and furnish useful information for teachers and students.

I just want to share with you a problem for fourth -- third- through fifth-graders that is in the NCTM book to give you an idea of what the image of effective mathematics and what is -- and what elementary kids ought to be able to do now.

The question is, what is the surface of each of these four towers, and as the towers get taller, how does the surface change? What this problem illustrates in our experience with this kind of problem-- Lora, who is a math teacher, reminds me that I should say the total area -- total surface area--

Did I get it right? (affirmative response)

I'm learning this-- I'm learning a lot of elementary math on this project, I assure you. (laughter)

What we're learning from problems like this, and how kids handle them, is that they can take data, as represented by these towers, and find patterns in this data. And from these patterns, they can make generalizations. So they can see the relationship between the height of the towers, as represented by the number of cubes, and the surface of the towers. And then they can take these generalizations and answer questions like, how many cubes would there be in a tower with a surface area of 242 units? I don't remember doing problems like that when I was in the fourth grade, so the bar is going up.

By the way, should you care, the pattern here -- I spent some time fiddling with it -- is that the surface -- the total area surface is four times the height plus two. So we won't make anybody work it out here.

Let me just say one other point about the blocks, because this is going to be important for some of the comments I make later. One of the ways kids learn these kinds of -- how to deal with these kinds of problems now is by getting manipulatives. And teachers talk a lot now about manipulatives. An example of manipulatives would be putting 15 or 20 blocks on a kid's desk and letting the kids build the towers and play with the towers and look at how the towers work and figure out the answers and look for the patterns. And, in fact, there's a lot of different kinds of manipulatives teachers have and kids have available to them now.

New Jersey's standards and assessments-- And I, too, am going to quote from the Achieve review of our standards. Just briefly, the math and language arts standards are a good starting point, says Achieve. But the standards need to be clarified. And they have some other suggestions about improving them.

Interestingly, Achieve is more complimentary of our assessments than of our standards. They say the math and language arts tests are challenging and measure important knowledge and skills. This isn't to say that New Jersey's tests aren't without problems. One of them that we've talked about a lot is time.

I do want to mention one validity issue. We've talked a lot -- or there's been a lot of mention of the language arts tests. The last data that we saw were that 55 percent of the students in the state passed -- were marked as proficient on the language arts test in the fourth grade, 66 percent were marked as proficient on the math, and 86 percent were marked as proficient on the science. That's quite a discrepancy. Science is not inherently harder than

language arts. And probably, if you look at what's been going on in classrooms, more time has been spent on teaching language arts. So there's always been sort of a question to me about why are the science scores so much higher than the language arts question. This is something we'll get back to.

So there's some background.

Now, we have been doing a study of the effects of math and science tests on fourth-grade teaching. We've been focusing on the ESPA. We have been doing -- focusing on math and science, because our research has been supported by the national science foundation. This is a three-year study. In the second-year study, we did a survey that included results from 287 fourth-grade teachers. The sample that we're working with was very representative by district factor groups -- so wealth -- and we had a very high response rate.

You can't use surveys alone, so we also did observations and interviews with about 60 fourth-grade teachers. And this is part of a larger three-year study. I will, at some points, refer to the first-year data, as well. So this is the background we're going to be speaking from when we talk about what we see going on in New Jersey.

And here's our first finding, which -- and we think it's important. ESPA is promoting more intellectually challenging instruction in math and science in the fourth grade. Our evidence from this comes in part from teacher interviews. Teachers say that they are using manipulatives more. So they're changing their methods of teaching. They're changing their questioning strategies. They're not just asking for the right answer. They're asking more open-ended questions. They're asking more how-did-you-figure-that-out questions. They're asking more why-does-that-make-sense questions. They're

asking kids to prove things more. This is what they're saying to us. They say they're having kids solve problems, and they're having students write more in math. So it's not just the matter of doing the work. These are all things that should get kids thinking behind the operations that they've always been learning to do. So that's the positive side.

Our second finding is that in spite of ESPA, teachers are not changing their instructional practice substantially. When we went in to observe, we did find that of the classrooms that we observed, 60 percent of them were using manipulatives in the classroom.

SENATOR MARTIN: Do you want to tell us what you mean by that?

DR. FIRESTONE: Examples of manipulatives is the blocks we were talking about before, where kids would get a set of blocks, and they could actually build the towers. There's something called Base 10 blocks where you can learn place value by getting some blocks that are a single square and some that are a row of 10 squares and some that are a cube of 100 squares. And by working with these things, they can see graphically what happens when you add 125 and 176 and physically see a representation of carrying and place values.

SENATOR MARTIN: Would computer simulations and--

DR. FIRESTONE: Well, this can be done with computers. This is often done with physical objects. We're seeing a lot of-- What we saw a lot of when we were observing was what we call candy math, where kids would get bags of M & M's or sets of M & M's, and they would have counting problems of one sort or another. And I guess a piece of the motivation was when they got done, after having done their counting task and their probability assessments

and things like that, they got to eat the manipulatives. (laughter) So it could be computers, but it could be much simpler and much cheaper than that.

SENATOR MARTIN: There's some presumption that manipulatives are a good thing.

DR. FIRESTONE: There is a presumption that manipulatives are a good thing, and the reason for that presumption is that when you-- Most of the math that I experienced when I was a fourth-grader was with numbers -- the written-out numbers. So we were doing problems like 175 plus 126. And some kids can understand what that looks like, and some kids can't. So when you have manipulatives along with sort of the normal numeric system, you can see -- you can actually see what you're representing with the numbers. So it makes it more concrete.

And as the example I tried to show with the towers illustrates-- It gives kids a chance to kind of explore the ideas, play with it, and learn things -- learn through play, I guess, is the way -- play is too strong a word -- but learn by exploring the concepts by using the activities so that they can kind of use the materials to kind of construct a representation and then an understanding of the idea.

So, yes, the presumption is that manipulatives are a good thing if they're used to allow kids to explore ideas, that is if they're not used as another kind of a -- another version of the conventional form of drill and practice -- 25 long division problems.

What we saw when we moved in and did the closer observation, though, was that when the manipulatives were used, the students used them to practice in the same old way 60 percent of the time. So that while

manipulatives could be good in theory -- if they were used for exploratory purposes-- Most of the time they're not being used that way. So there's a potential that's often not being realized.

Also, most of the assignments asked kids to practice algorithms like how to do long division, like how to multiply fractions, or things like that. They were not getting into the issues of the principles behind the practices that you have to follow or the procedures that you have to follow to get into understanding mathematics more deeply.

And finally, teachers only asked for short answers. Although they said they were asking for more explanations, when we sat down and watched, they tended to say, "Okay, what's the answer?" And Johnny would say, "25." And the teacher would say, "Right. Okay, Jill, what's the next answer?"

So while teachers are starting to think about changing their practice in a way that would lead to more challenging instruction and kids having a deeper understanding, this tends to be something that's kind of grafted onto the old ways of teaching math, which isn't going to get them into things -- get children into a deeper understanding of math, and similarly with science. So that's finding No. 2.

Finding No. 3 concerns content change. And here, the patterns are a little bit different in math and science. This chart shows change in the numbers of teachers spending a lot of time on certain topics. And the red column is how much -- the percent of teachers who spent a lot of time on the topic in 1999 in the first year of the study, and the gray column is the amount of time they spent in 2000.

So what you see is a clear drop in the amount of time spent on whole number operations -- addition, subtraction, multiplication, division, that kind of thing -- and you see more time being spent on fraction operations, on statistics, on patterns and functions. These are topics that are part of the ESPA. These are topics that are reflected in the standards. And particularly topics like statistics and patterns and functions are things that were not typically part of the fourth-grade curriculum before the core curriculum content standards came into being.

So what we see here is the kind of change that you would expect to find from the tests being put into place and the standards being put into place.

The story in science is a little bit different. We looked at three topic areas here: investigative skills, structure and properties of matter, and states of matter. And what we see from '99 to 2000 is a reduction in the amount of time spent on these topics.

Now, what's really interesting about that is when we surveyed the teachers in '99 and we asked them are you increasing or decreasing the amount of time you're spending on these and other topics, they were saying they were increasing. In '99, when we just asked this question about what are you increasing and decreasing, there were a lot of increases. In 2000, there were a lot of decreases. And we think this is a funny kind of teaching to the test. We think this is a response to the fact that the proficiency levels are such that more kids are passing the science test than they are passing the language arts test.

So when there was the threat of the new science test, there was a lot of pressure for teachers to learn more science and spend more time on science. But as they saw that, gee, whatever they were doing was fine in science, and

they're having a lot of problems in language arts, suddenly they appear to be shifting more of their attention to language arts.

We can't comment on the language arts side of that, because we didn't ask questions about language arts. But that should give you an idea of what we see happening in the field and our guess as to why that is.

SENATOR MARTIN: This portion of it is based upon their responses? This isn't based upon observations of time spent?

DR. FIRESTONE: This is not based on observations, this is their response is right. It's very hard. I mean, we're talking about our metric here is devoting more than 11 lessons per year to a topic. We would have had to spend hours and hours in classrooms.

SENATOR MARTIN: I'm just trying to understand how it's done.

DR. FIRESTONE: Right.

SENATOR MARTIN: And it would be very hard to--

DR. FIRESTONE: Right.

SENATOR MARTIN: And it would raise a whole host of -- if you were doing some kind of visual eavesdropping or something like that. But this is based upon their indications. We're not clear whether the time frames you suggested-- They may still have, supposedly, an X block of instruction for science, which is supposed to devote hour -- between Hour 1 and 2 to science. And if it's not-- If they're reducing the time here-- You're just assuming there's a-- They're now using the difference in time to spend in, perhaps, another area, which is more problematic.

DR. FIRESTONE: Yeah, that's our assumption. And we would be stronger on that point if we had the information for language arts.

And I think you're right, too, in saying we need to know the block scheduling now -- or the scheduling system. Elementary schools are more flexible than high schools. When we asked teachers how much time they were spending on subject areas -- math, science, also recess, music, those kinds of things -- we were not finding big changes. But there does seem to be a reduction in the amount of time going into--

When we asked them about a whole set and we asked them about 17 science topics across a whole large number of them, what they seem to be saying is it's being reduced; less time is being spent on this.

SENATOR MARTIN: Senator.

SENATOR BAER: I wanted to ask, have you not only gone into the best way of the teachers teaching to the youngsters, but the best way to teach to the teachers to teach this way and the best way for the administrators to monitor and follow through to see that as much of this is incorporated?

DR. FIRESTONE: We asked a set of questions about teaching practices teachers were using. I think this is where you're going. We did ask a set of questions about teaching to the test and what they were doing in response to the test, per se. And what we found was that they were doing a variety of things. Some of the things they were doing were teaching test mechanics, which would be things like how do you bubble in those little forms so that you get your name on things right, you make sure that your answer is recorded right. Test mechanics would also include things like pacing yourself through the test.

We asked if people were using commercially available test preparation materials. We asked if they were using rubrics, which is another one of these new developments that comes up largely in language arts, but it's sort

of a scoring system. And by having kids kind of discuss and learn about the scoring system, they usually kind of understand the structure or the discipline and the thinking style and discipline better.

I'm not sure I made myself clear. Let me try again.

SENATOR MARTIN: Byron, if you're going to -- come back to your seat.

MR. CANNON: We don't have a microphone there for the hearing unit.

SENATOR BAER: The button was red, and I thought it was picking me up.

SENATOR MARTIN: That's for the transcribe-- We're talking about transcribing it.

SENATOR BAER: Oh, okay.

I'm not sure I made myself clear. Maybe I did. But the things you're describing seem very interesting as effective ways of teaching if they're adopted. So my question has to go to have you not only gone into that, but how to get teachers to do this more. What is the best way of teaching the teachers to do these things? What is the best way of getting administrators to follow through and see if these things are being done? Is that within the scope of your work?

DR. FIRESTONE: That is. If I can get through one point-- We don't--

Let me say, right now, we don't have information on what principals should be doing. We do have some information on what teachers

should be doing. And if I can get through one more point, I can get to that, if that's all right.

SENATOR BAER: Sure.

DR. FIRESTONE: Okay. Are we clear on the science part?
(affirmative responses)

Another finding is that teaching to the test is not promoting inquiry in urban schools. And we had a scale where we asked teachers questions about how often they're doing things -- certain kinds of things to prepare for the test. And it's better to look at this chart, although the key is a little small there.

Three questions we asked teachers were: Are you using rubrics, or are you having kids use rubrics to score their own work, which should help them understand things more; are you teaching test mechanics, which is kind of--

One of the things-- We spent a lot of time talking about what does teaching to the test mean. And I think the kind of most mundane form in the sense of teaching to the test is the kind of teaching test mechanics kind of thing.

And are you using commercial materials, which would kind of be very much focused on drilling on the test.

And if you look at this chart, what it shows is, the sets of columns from left to right are district factor groups. So the set of columns on the left is the poorest school districts. The set of columns on the right is the richest school districts. And what we found was the most teaching of test mechanics, the most use of commercial materials, those kinds of things, happened in the poorest school districts.

That didn't happen with using rubrics, which we think is a more intellectually challenging form of test preparation, which we found kind of

equally distributed between the wealthiest and the poorest school districts. So what we see is more of a kind of drill focused instruction in the poorest districts. We don't necessarily see less of the more kind of intellectually challenging focus instruction in the poorest districts.

So it's not doing more encouraging of intellectually challenging instruction in the poorest districts, but it is encouraging the drill.

SENATOR TURNER: They don't seem to be benefiting from that. Is that-- Even though they're getting this drill instruction, their scores are not improving as a result of it.

DR. FIRESTONE: It does suggest that more drill, in and of itself, is not going to raise the scores, but I think you'd have to be doing more analysis than we did linking the test data to our kind of teaching data to be sure of that. But it's suggestive.

SENATOR TURNER: Would you say that -- or I don't know if you would know-- Would you say that the tests are culturally biased in any way?

DR. FIRESTONE: I'm not the person to answer that. I would not think so. I think if the tests reflect-- My guess is, if the tests accurately reflect what adults need to survive -- to function effectively as adults in society, then it seems to me that those are useful standards to apply to all kids of all races and cultures, because they're going to be coping with the same issues in the workplace. That would be my take on it. My take would be, it's important to give the kids the skills to pass the test, not to change the test so that everybody could pass them.

Let me move on to a point that speaks to Senator Baer's concern.

Change in practice-- What we were saying is, teachers are starting to change their practice, but they haven't moved as far as they ought to -- is being limited, because teachers' learning opportunities are limited.

We asked a lot of questions about access to professional development. And here is one that we think is an important set of findings. The research now is suggesting that the one-shot, one-day professional development workshop, which has been the typical pattern, doesn't really help teachers develop the skills and knowledge; that they need to spend extended periods of time working on the same topic in the same area and linking it to their classroom practice to get better at it.

So we asked teachers whether they spent more than two days on professional development on a certain topic in a given year. And this chart shows the results. And it's a kind of good news -- in our view -- bad news kind of a chart in that it shows more teachers saying they spent extended periods of time in professional development on the same topic in the same year. So from '98 to '99, the results in content and instruction related to science went up from 22 percent in '98-'99 to 31 percent in '99-'00, and similarly with content and instruction in math.

We tried to break this out to differentiate between professional development that really focused on the concepts and the ideas and how to teach them as opposed to test prep professional development.

We're not sure teachers always made that distinction really clearly, but that was what our effort was. And what you see is, in all of those areas, there's an increase from '98-'99 to '99-2000. But we think teachers still need

more extended periods of time to develop the skills they're going to need, and we say that partly based on the kind of classroom lessons we saw.

One last thing on teaching to the test before we move on-- One of my colleagues, Roberta Schorr, who's at Rutgers Newark, spent a lot of time looking at the observation data. And as she sorted through it, it looked to her like, and I think we have some quantitative data for this, as well, the teachers who broke out teaching to the test and made that a separate activity tended to be the teachers who didn't seem to understand the subjects as well. Teachers who understood the subject could integrate teaching the basic concepts of math and science with preparing kids for the test, so it wasn't this kind of dichotomy. And that's another reason why we think professional development is a key issue.

So let me jump from that to a little more speculative exercise in what we think criteria are for a good testing program. First of all, tests should reflect important content and skills, which should come about with good alignment with State standards. Second, tests should be good measures. They should be valid and reliable. Third, tests should provide information that's useful to teachers and parents. Fourth, tests should illustrate good instruction. And the parts of the test in New Jersey that do that are the open-ended questions.

And I want to say I think a lot of the movement we saw where we said teachers are starting to try new things-- They told us that was in response to the open-ended questions.

Finally, tests should not cost too much in terms of time or money.

The problem is, the Department and the Legislature has a number

of trade-offs to consider. It's very important to have open-ended items, because they do challenge conventional practice.

The Bush administration is suggesting that annual testing is important. We think it is useful, because it shows growth. One thing that might help equalize the difference between rich and poor school districts is in looking at year-to-year growth. Now what we're doing is, we're looking at test scores at one point in time, and it's hard to sort out the effects of what kids have learned in school from the effects of what kids bring into school with them. If you look at growth, you would look at the effects of what the schools were teaching better.

Third, testing more subjects and topics ensures that educators will attend to them. There's a risk that what doesn't get tested won't get taught, but the more of each of these things that you do, the more it's going to cost. It's more expensive to score open-ended items than it is to score multiple-choice items and the more it's going to take time away from instruction.

Now, we only test in third, fourth, a little bit in fifth, and in eighth grade. We may find ourselves testing in more grades. So there is a set of trade-offs that have to be considered.

Our recommendations of what we think are important-- First of all, and we've mentioned this to the Department, we think it's important to coordinate proficiencies across subjects for the reasons that we showed. We don't think the amount of time spent on science should be reduced because the proficiency level has been set higher -- set as easier to pass in science.

Now, we've talked this over with the Department, and we recognize there's a dilemma here. This whole business of what we're using is criterion

reference test, which means we determine what is proficient. We aren't talking about who's above average, who's below average. We're talking about who is scoring at a proficient level, but there's a certain amount of judgment involved in determining what constitutes proficient. And the testing community -- the testing professionals around the country don't have a lot of experience in figuring out how to kind of equalize proficiency across subject areas. So we think it's important to work on, but we recognize it's a tough nut to crack.

Second of all, we think it's very important to stay with the open-ended items, because, as we say, that's been the lever for improvement in New Jersey.

SENATOR MARTIN: Just on the point before you got to that, what you seem to be suggesting is, at least -- and what the data sort of suggests is that the proficiency level that we set at science is much lower than the language. At least that's one way to easily account for the fact that so many fewer kids are passing the language skills test as opposed to the math.

DR. FIRESTONE: That's a reasonable interpretation.

SENATOR MARTIN: There are others, obviously.

DR. FIRESTONE: There are others, obviously, right. But I think that needs to be looked into. And so far, as I understand it, the procedures for setting proficiencies has been to work in each area independently.

SENATOR MARTIN: I suppose-- I would think that perhaps some other studies may suggest that science is not instinctively easier to do than language skills. If anything, it, at least-- I think the common general sense was that it was at least as hard, or harder, at least, for many people to do science as it was to do language skills.

DR. FIRESTONE: You would think. And it also depends at what level of science you're doing. I mean, science-- I mean, all curricular topics can be done in such a way that they're easy or they can be done in such a way that they're hard. So that's a piece of it, too.

The other piece, I would say, is we don't have before-ESPA time estimates. But our guess is, from talking to most people, that not a lot of time was being spent teaching science before. So it's strange to us that so many more people are passing in that area than are passing in language arts.

SENATOR TURNER: But does that-- Well, what answers the question as to why the urban students are closer to their counterparts in terms of the scores on the science test?

DR. FIRESTONE: What answers that?

SENATOR TURNER: Why?

DR. FIRESTONE: I don't really have a good answer to that.

SENATOR MARTIN: But if it's just a pass test, it doesn't necessarily say that. It just means that there's a higher percentage of them passing, but they're not measuring--

SENATOR TURNER: I would say it's proficiency, isn't it? They use the proficiency--

DR. FIRESTONE: Yeah. To pass-- A technical process has been gone through to set a level which is determined to be proficient by working with experts in the field who look at the test items. So passing so many items or certain items constitutes proficiency. And that is done separately in language arts, math, and science.

SENATOR MARTIN: But my point is, suppose you--

SENATOR TURNER: But there's a larger gap. There's such a wider gap in terms of language arts. But then when it comes to the science area, that gap is narrowed a lot.

LORA MONFILS: That's basically what we're asking.

DR. FIRESTONE: That's part of our question. I would guess, and Lora might be able to help me out here, some of that may just be because, overall, more people are passing in science.

SENATOR MARTIN: The gap could be equally different. Suppose the cutoff level to use SAT scores was 500. So maybe everybody is reaching 500, because 500 is-- The questions now allow for more -- 500. But the top performing students, on average -- let's say suburban schools could be getting an average of 700. We're just deciding how many have passed. But it could be, if the gap is here in language over math and science, it's there. The gap could still be the same. It's just that the lower level has reached what they've determined proficient. Whether that is proficiency or not is an open question, as well, since it's relatively new. But we don't know. It may be close or it may be--

SENATOR TURNER: Well, that's why I'm questioning the reliability and validity of the test.

SENATOR MARTIN: We're not measuring the differentials, we're just measuring whether you passed or not.

DR. FIRESTONE: Let me say that's a question that our study was not really designed to address. So what I would give you at this point would be more than the normal amount of guesswork at this point. It might be something to explore with the Department. It might be something that-- It might be that

somebody could do some further analyses of the existing data and throw some new light on that.

SENATOR MARTIN: Theoretically, the data -- at least some of the data should be able to more accurately answer that question, because above the proficiency level, there is still-- There's a range, one would hope, large enough to be able to ascertain whether those in the urban schools that had passed were getting the same relative scores of those who were passing in the suburban districts or whether there was still that gap.

DR. FIRESTONE: Conceivably, depending on what-- I mean, there's a theoretical range. How well that was tapped in the assessment, I can't really say. So again, that's something that would be better discussed, I think, with the Department than something we can answer directly.

SENATOR MARTIN: But it's an important issue.

SENATOR TURNER: No, I was going back to what you said. Ordinarily, you think of someone who does very well in science -- would be, pretty much, above average in language arts in terms of testing. Generally, if you--

SENATOR MARTIN: I mean, at least -- sort of the general knowledge is that science is harder than language skills -- language arts.

SENATOR TURNER: Right. And when you take a test, in most cases, it's testing your language skills. You have to be able to read and comprehend.

DR. FIRESTONE: Yes.

SENATOR MARTIN: One might also think that a Department fearful of having bad scores might have, the first time around, set the bar at science and math very low and then everybody -- or so many people passed.

SENATOR TURNER: That could be.

SENATOR MARTIN: That's always another inference.

DR. FIRESTONE: I have to say-- I mean, that's a good hypothesis. I have to say, though, as I understand this, these scores were set with a fair amount of input from outside subject matter judges. So I think it's important to look at the procedure.

What I'm suggesting at this point is, it's important to look at the procedure for setting those proficiency levels.

SENATOR MARTIN: Well, the other part of it is they may well, in terms of validity, measure the proficiency for a certain body of knowledge which was tested. The question is, if it's so easily achieved, is it a level that we want to be satisfied, especially when measured against other countries where we know that their science and math skills, apparently, are being -- are at a higher threshold of knowledge.

But why don't we proceed here.

DR. FIRESTONE: Okay.

Another thought that we have is consider adding subjects over time -- testing kids in more subjects in the eighth grade than in fourth grade, and maybe at the high school level having end-of-course tests that are sort of integrated with the course materials of the high school rather than a single kind of test -- single HSPT or HSPA type test.

Also, and I'm not sure what's going to be legal from the Federal perspective soon, consider testing subjects in alternate years so that every subject is not taught in every year if there is a concern about too much time.

So these are some possibilities to think about.

But I want to change the subject and share this with you and suggest that you don't grow cows by weighing them, by which I mean, kids don't learn by taking tests. What testing needs to be thought of is part of a broader systemic approach to improving education in the State. And the State is doing a number of things to improve education now.

In the urban districts, there's the early childhood programs, there's whole school reform. But outside the urban districts, efforts to improve instruction, very considerably in quality and sophistication-- Small school districts are at a disadvantage, because they typically don't have the central office staff of specialists who can help teachers learn the content in these areas that are tested. Poor, non-*Abbott* districts are also at a disadvantage.

So we think that the State ought to think about a three-part program, part of which includes professional development for current teachers. We have the 100-hours requirement now, but we don't know yet how it's working. And we believe that most teachers need to know more about the content they're teaching, how kids make sense of it, and how to engage kids in learning that content. So we think there's a need for more professional development.

Think about materials. A recent report found that no middle school textbooks met national science standards. New Jersey is not a textbook adoption state like Texas and California. And those states don't share our

standards. They certainly don't share our assessments. So they're not likely to be creating a market for the kinds of materials that would be well aligned with our standards and our assessments.

And think about teacher certification, which other people at Rutgers can say more about than I can.

But I think there needs to be movement on all of these fronts if we're going to see improved learning in New Jersey to go along with the standards and the assessments.

I do want to mention that in some ways, higher education may be an underused resource. The traditional role for higher education has been preservice education only. We are now, I would say, from my personal experience the last two or three years, entering kind of in an era of good feeling with -- between Rutgers, at least, and the Department that has allowed us to provide a lot more service for the State.

And some of the other things that could be done would be provide a continuing intensive professional development. With another hat, I direct Rutgers' Eisenhower professional development program. And there is -- national assessments have been done. Eisenhower is a federally funded program to provide professional development to teachers in math and science. And it's got one piece that gives money to schools and districts, and one piece that gives money to institutions of higher education.

The evidence is the funding that goes for the institutions of higher education leads to greater change in teaching practice. Now, some of that is because the programs are more intensive, but there could be a large-- It suggests

that a larger role in continuing professional development for institutions of higher education would be in order.

Another possibility is kind of quick reference services for educators providing Web sites of various sorts. Rutgers is helping with the New Jersey PEP.

A third possibility would be informing parents. A lot of the language about the new standards and new assessments is foreign to a lot of parents, because things have changed a lot from when they were in school. And programs focused on helping parents understand and be able to help their kids can be very beneficial.

Finally, there's a number of questions that have come up today that can be addressed through program evaluations, and the institutions of higher education in the state have a lot of capacity for doing that kind of thing.

Finally, as you move into an implementation strategy for not only testing but improving education in the State, a few thoughts. One, more three-way planning between the State, the districts, and the institutions of higher education-- A lot of the people in the districts, I feel -- I talked to feel like mandates are being laid on them. They can make very useful input. I think the institutions of higher education can make very useful input.

Second, as with this hearing, be very public. The standards and assessments are very confusing to people. The more the public comes to understand these things, I think, the more comfortable they'll be with it.

Third, be in the mode of experiment and assess. We're constantly in the process of developing and improving the State assessments. We should be in the process of developing and improving strategies for teacher professional

development and teacher education on all fronts. We're not going to get it right the first time, but if we keep looking at what we're doing, we'll get better.

And finally, stay the course, but fix the details. I say this in an election year, because many states have a history of making radical changes in direction when there's a new administration. I think New Jersey is developing a strategy around its standards and assessments that is useful. It's only been in place for five years. And I've talked to so many teachers in New Jersey and in other states who say, "Well, you know, we can wait this one out, because this, too, shall pass. Policy will change." So whatever we choose to do, I think we should stay with it long enough so that it can have some benefits.

And with that, let me stop and answer questions.

SENATOR MARTIN: Senator Turner.

DR. FIRESTONE: Did everybody get their licks in? (laughter)

SENATOR TURNER: You were very thorough.

I don't know. Do you have any way of knowing, in terms of the urban districts, the mobility of the student -- how it effects the teaching or the assessment of that student? For instance, one teacher told me that when she starts her class in September, by the time June rolls around, she has -- 45 percent of those students are not there. And she has new students coming in from other schools. So I would think that would have an impact on the test results. You're not really testing or holding that school accountable for the test of that particular -- or those students, because they're not the same students that have been there all year.

SENATOR MARTIN: I just want to say, there have been national studies. The best ones I've seen that deal with what -- that problem, which is

real, especially in the urban areas, is called Value Added Testing. The University of Tennessee started some studies some years ago in which they were able to look at students at the beginning of a school year and at the end of a school year. But it does require some form of testing. But it's a way--

It's also a way of measuring teacher performance, as well, because instead of penalizing teachers because they may have students at different learning levels when they come in, they're able -- they were able to make assessments of how effective the teacher has been over a period of time. Absent that, you run into problems, such as you have students with differences in preparation. And you also have that huge mobility issue, which is trying to measure across districts and even schools in New Jersey -- is very problematic, especially in the urban areas. Because it's simply not fair -- where you have as much as 30 to 35 percent mobility in some of -- for example, Newark schools -- to ask a teacher to take a student who wasn't her student and then assign her responsibility if the student doesn't do as well as she or others would like.

But I don't know how you'd capture that without some form of testing, but we'll let Dr. Firestone talk about that.

DR. FIRESTONE: I think we're in a lot of agreement here about the fairness of holding teachers accountable for the progress of kids they haven't taught all year long. I will add to this that I believe on the report cards there is mobility data. So one can see school by school how much turnover there is. And one of the analyses we intended to do as part of this project is look at the variation and mobility by SES and by district factor group and see how that ties in, but that's an analysis we haven't done yet.

SENATOR MARTIN: The other comment I would add-- At least some of my experience in dealing with teacher preparation-- I've had different proposals. And I recognize-- I think there is enormous benefits of using the higher education institutions, especially Rutgers. But there has been some -- how do I say this -- fair amount of criticism in the past, or at least concern, in that higher ed hasn't done as good a job as it could in the confidence level, at least maybe not in this era of good feelings that you've suggested, but that local school districts and principals and certain school administrators are more inclined to want to do the preparation at this local school district, as opposed to relying on certain types of continuing education and performance instruction at universities based upon their, at least, experience from yester-year about the skills that were being taught at our institutions.

What I'm suggesting is, if there was a better -- I don't know whether it's a problem or whether today it's a marketing issue -- but I think institutions like Rutgers that are suggesting that they can really provide services in this area of teacher performance have to convince -- have some convincing to do, because I think there have been some concerns that, at least up to now, the data -- that hasn't always been the case. And there's been a lot of courses and a lot of-- Just like those one-day ones that you dismissed, and I think rightly so-- There's also been some instruction that institutions of higher learning in New Jersey -- that at least teachers and principals have not felt have provided as much as -- service as they should and could have.

So I don't think I'm telling you something new. It's maybe not the most greatest news, but hopefully the new word will spread that Rutgers really has programs that are able to help teachers with the core curriculum content

standards, as well as our other schools like Kean University and Montclair State and so forth.

DR. FIRESTONE: Let me unpack that a little bit. I didn't mean this just as a plug for Rutgers, although my experiences with Rutgers -- there's a number of institutions in the State who can help. And I think within Rutgers and across the state there has been variation in the service that is offered. And I suspect we will also find, as we dig into this, that the one-semester course at the university is not necessarily the best model for helping teachers learn how to improve their practice, although it has the advantage over the one-day workshop, if it's structured right, and they aren't all -- if it's structured right, of giving teachers assignments that they can take back to their classrooms and try out and come back and debrief and debug.

So I think there's some work to be done on getting institutions of higher education and K-12 institutions working effectively together. You're right, it's partly a sales issue, it's partly helping people on both sides understand the standards, and helping people in schools sometimes understand the challenge of the standards and helping people in the universities get their feet on the ground about the realities of schooling. But there's a lot of potential there that I think we could be taking advantage of.

SENATOR MARTIN: I agree.

And there is some new initiatives. I know, at least on the Rutgers Newark campus -- that urban initiative that's sort of multidepartmentally based, but it has a huge education component where it connects directly to the *Abbott* school district and specifically the Newark school district -- is, I think, encouraging -- because looking at sort of a more holistic thematic approach--

Thank you very much. I appreciate it.

DR. FIRESTONE: Thank you.

SENATOR TURNER: Doctor, I want to ask this question, please, before you leave.

Doctor Firestone, have you found during your studies that there's any correlation between test scores of the student and the education level of the parents?

DR. FIRESTONE: I need to be clear that in this research, what we have been looking at is teachers. I think there is data that shows that kind of relationship. But that's in the literature. A number of people can comment on it. I think a number of us have seen it. It's interpreted in different ways at different places, but that's not been what we've been focusing on.

SENATOR TURNER: I don't think--

DR. FIRESTONE: What we've been focusing on is trying to figure out what's going on in schools, with an eye of figuring out what can we do in school to improve education.

SENATOR TURNER: So you're just looking at the isolated area of the school.

DR. FIRESTONE: Right, our focus has--

SENATOR TURNER: But I think most of--

DR. FIRESTONE: --been working out from teachers.

SENATOR TURNER: But I think most of it has to do with what the student is bringing from home, in terms of the home environment, and the kind of nurturing and education they get in the home and whether or not that

student is being helped by the parent in terms of education at home, and also, exposure.

DR. FIRESTONE: I wouldn't disagree with that. I think, in a sense, what we're doing is trying to take that as given and therefore say, so what can the schools do in that context to do their job better? And that's been what our work has been focusing on.

SENATOR TURNER: Okay.

SENATOR MARTIN: Thank you.

DR. FIRESTONE: Thank you.

SENATOR MARTIN: Do we have representatives from the New Jersey Chamber of Commerce and Local Initiatives -- Prudential Insurance Company? Is that Mary O'Malley, Dana Egreczky?

We would appreciate any manipulatives for the Committee to help us understand this process. (laughter)

DR. FIRESTONE: I'll bring M & M's next time. (laughter)

SENATOR MARTIN: I like the edible ones.

Thank you.

MARY O'MALLEY: Yes, we're a tag team.

Mr. Chairman and Committee members, I appreciate the opportunity to speak with you this afternoon.

My name is Mary O'Malley, and I'm a Vice President in Prudential's Community Resources Department. In my job, I'm head of community affairs for Prudential and part of a team of Prudential employees who are assigned to work on public education reform.

As part of that team--

SENATOR MARTIN: How do you get assigned to that? Is that a voluntary--

MS. O'MALLEY: No, that's part of my job. I get paid to do that.

SENATOR MARTIN: No, but you said this department-- Is it a special subdivision of Prudential, or is it--

MS. O'MALLEY: Yes.

SENATOR MARTIN: It's not volunteers like doing a community service project, it is--

MS. O'MALLEY: No, no, we have -- we mobilize volunteers for Prudential. There are three components of this department. It's corporate philanthropy. Prudential gives away \$25 million a year through the Prudential Foundation. Then there's a social investment program that's manned by investment professionals, where we make socially responsible investments in communities to organizations that might not otherwise be able to attain financing. And then there's the unit that I had that's called, traditionally, community affairs, which is volunteerism and executive involvement in community involvement in government relations and things of that nature.

As part of that team, my responsibilities include supporting Prudential's Chairman Art Ryan's public education initiatives; running New Jersey United for Higher School Standards; heading the executive committee of the Business Coalition for Educational Excellence. I also led the teams that conducted New Jersey's two education summits. In addition, I helped plan the last National Education Summit and am helping develop the one that will take place later this year.

In my testimony this afternoon, I'll focus on two pieces of research that Prudential has funded. If you can stand it one more time, I'm going to talk about the Achieve study that looked at the core curriculum content standards and assessments. And I'll talk about it in slightly greater detail than Drs. Firestone and Gagliardi did.

The second bit of research is a straw poll conducted in a series of town hall meetings held throughout the state by New Jersey United for Higher School Standards.

Here is the position I will support with this research. The ESPA and the GEPA are strong learning and assessment tools. They can play an important role in driving learning and in helping improve our public education system.

Parents, educators, and the general public have some justifiable concerns about the administration of the tests. But their most frequently articulated reservations center around need for greater clarity in the standards and around the adequacy of local resources, particularly professional development. They also speak to the need for increased teacher professional development opportunities to help students meet the new standards.

Let me start with the highlight of the earlier research. You heard about Achieve. It's a nationally respected, nonprofit organization that is best known for convening the 1999 Education Summit. Led by governors and CEOs who represent a broad political spectrum, Achieve supports the implementation of standards-based education through research, diagnostic benchmarking, convenings, and partnerships.

A grant from the Prudential Foundation enabled Achieve to evaluate New Jersey's language arts and mathematics standards and assessments. Nationally recognized experts compared our system to standards-based systems that are recognized among the strongest in the world, including California, Massachusetts, Arizona, and Japan.

In brief, here are Achieve's findings. New Jersey's language arts literacy and mathematics standards have some important strengths, but they must be made clearer and more specific if they are to dramatically raise student achievement. The standards do not clearly communicate what students should know and do to succeed on the State assessments.

But, New Jersey's assessments in language arts literacy and mathematics stand out among the nine states that Achieve had reviewed at the time of the report. With three key exceptions, the tests are challenging and measure important knowledge and skills. The items on the tests are reasonable for the grade levels tested, and the items tap both basic knowledge and skills, as well as higher level skills. Achieve's concerns about the assessments addressed the HSPT mathematics test, which is seen as insufficiently rigorous; the length of the language arts literacy tests; and the scoring of its open-ended items.

Achieve's recommendations for going forward are that the Department of Education should build on the strengths of the State assessments and revise the standards to provide explicit guidance to educators and the public about what all students are expected to learn. Speaking specifically about the assessments, Achieve noted, "With relatively minor revisions, the assessments can be made stronger to ensure that they provide solid information about student progress toward the standards."

The Achieve report offers both compelling and interesting support of the assessments. The findings are compelling because, frankly, they were the result of an extensive study by nationally recognized experts conducted by an independent organization. They are interesting because they were, at the time, the reverse of what was being discussed in the press.

This report was released in December by Achieve and Governor Whitman. It's been widely distributed throughout the State and has been positively received because of the rigor and the quality of the benchmarking. In fact, this report stimulated Merck to commission a similar review of the science standards and assessments, which is in progress and which will answer some of the questions raised today on the science assessments.

At the same time that the Achieve study was under way, New Jersey United for Higher School Standards was conducting a straw poll of public opinion about the standards and assessments system.

New Jersey United was formed in the fall of 1999 at the time that the first ESPA and GEPA scores were being released. Its founding co-chairs were Art Ryan, Chairman and CEO of Prudential; Al Cade, then chair of the Commission of Higher Education; and Bob Bonazzi, head of the New Jersey Education Association.

The organization was created in order to allow New Jersey citizens to have a forum to discuss their concerns about the standards-based education system. In meetings, participants hear a brief overview of the system and then have an opportunity to raise questions or comment on areas of particular concern. We've conducted these meetings in 12 legislative districts so far, and roughly 1000 people have participated. At every meeting, we have members of

the local school system present, along with representatives of the Department of Education and the business community.

We've recorded the proceedings of each meeting and transcribed them. They are available on the New Jersey United Web site. In about a month, we'll be releasing a summary of our findings and a review of New Jersey United's first year of operations.

For the purposes of this hearing, I'd like to provide you with some preliminary highlights of what we've uncovered in our work.

What we haven't found was broad public opposition to the ESPA and GEPA. What we did find was a lot of expressed confusion about the standards and assessments and what they meant for individual students. Participants raised concerns about the assessments and some of the specific weaknesses noted in the Achieve report. They also strongly emphasized the need for increased professional development opportunities for teachers, as Dr. Firestone talked about earlier. The teachers indicate to us that they need more information on specifics in the assessments, and all of the participants strongly support a technology-rich learning environment.

Our New Jersey United straw poll findings, like the Achieve report, are interesting and compelling. Frankly, we expected to hear a lot more adamant opposition to the assessments. It is interesting to note that the concerns raised about the assessments largely recognize that the problems are fixable and that our standards-based system is worth repairing. What is compelling about our findings is that they suggest that the standards-based system has mobilized education stakeholders to want to get more involved in

education to have the knowledge and tools that they can use to help their child and improve their schools.

On the strength of what we've uncovered at these community meetings, New Jersey United is commissioning a survey of public opinion on standards, assessments, and accountability in public education. We expect to release that study in the fall.

Let me restate our original position. A growing body of research shows that the ESPA and the GEPA are strong tools that can drive greater learning in our schools. The acknowledged weaknesses in them can be fixed, especially if the upcoming standards review process results in clearer and more specific standards that communicate, with greater precision, what students need to learn. Our next steps should be to involve a broad range -- and, Dr. Firestone, I'd say more than three -- a broad range of education stakeholders in addressing the tests' problems and in ensuring that our students and teachers have access to the resources they need to meaningfully implement the standards and assessments in all of our state's schools.

I'll be happy to take your questions.

SENATOR MARTIN: Why don't we hear from our representative from the Chamber of Commerce.

DANA EGRECKY: Before I start my formal comments, I should mention that when business comes to the table to talk about education, we're not saying that we have any expertise in education, but, in fact, we're standing on years and years of experience and systems implementation and accountability and continuous improvement and, perhaps most importantly, the ability to attract, retain, and train a high quality workforce.

My name is Dana Egreczky, and I am Vice President of Workforce Development for the New Jersey Chamber of Commerce. And the Chamber represents over 2000 of New Jersey's most prestigious corporations. And together with regional and local chambers, we represent over 12,000 large and small businesses.

For the last several years, our employer constituents have been telling us loud and clear that their number one concern is the lack of a qualified workforce. National surveys indicated that 94 percent of employers admit that their inability to hire qualified workers has negatively impacted their economic growth.

Therefore, the State Chamber established Workforce Development as a key component of its advocacy programs, and its initiatives in that area extend into the K-16 educational sector. The State Chamber relies on the Business Coalition for Educational Excellence, which is a grassroots, business-led effort, to guide its programmatic and policy activities in the education reform arena.

Recently, the Business Coalition for Educational Excellence released a position paper called *Getting Down to the Business of Learning*. It outlines the business agenda in education reform in three critical areas: standards and assessments, teacher quality, and accountability. It also defines our commitment, the business commitment, to staying the course. The New Jersey Chamber of Commerce Board of Directors unanimously adopted this document at its April meeting. The Chamber's Board consists of over 50 presidents and CEOs representing a broad array of our member corporations. I am pleased to provide you with a copy of that position paper today -- that's the bound

document you got with this testimony -- and to distill from that some of the critical points made in it regarding the State's assessment from the business perspective.

The business community strongly supports a standards-based educational system inclusive of high-quality assessments that will measure student achievement against those standards. And we believe that all New Jersey students should receive the educational, parental, and community support they need to achieve New Jersey's core curriculum content standards.

But we acknowledge that New Jersey's assessment system has been the subject of considerable discussion since its adoption. Many organizations and individuals have commented on the State's tests in a variety of areas, including the grading of the instruments, the quality of test questions, the length of the tests, and the frequency with which they are administered.

We believe that the frequency and duration of any statewide assessment should be determined by the need for data, the ability of the assessment to promote continuous improvement, and research findings regarding effective statewide assessment practices. A critical component in determining the frequency of statewide assessments is the capacity of the system to generate consistently high-quality test instruments that measure higher order cognitive skills that are worth teaching to, since it is inevitable that mandated assessments will modify classroom practice. And I should mention to you that I personally spent 16 years of my life as a public school science teacher here in New Jersey and in Pennsylvania in middle and high school. And I do know that mandated assessments really make teachers sit up and pay attention.

An independent review of the State's assessments in language arts and mathematics by Achieve, Inc., as you heard, indicates that they are high-quality tests and are academically rigorous. However, it is our observation that New Jersey's system of assessments requires more development before it can provide important, reliable data. Once these tests are capable of generating meaningful information that guides districts in their continuous efforts to improve student achievement, then the State may wish to consider testing more frequently with shorter assessment and with a wider variety of instruments.

To ensure that the best possible assessments are developed, the Department of Education should incorporate a more inclusive audience of educational practitioners, content matter specialists, and business volunteers in the assessment development process. Like standards, assessments should be benchmarked against tests from other states and countries that are best in class. An external, unbiased evaluation of New Jersey's statewide assessments by an appropriate organization should be routinely accomplished and driven by legislative mandate.

Our final observation relates to the allocation of time, resources, and student attention divided among seven curricular areas and additional workforce readiness skills. Business leaders have learned that it is difficult for any organization to be consistently excellent in a boundless array of products or services. We are concerned that having full-fledged standards, curriculum frameworks, and paper and pencil assessments in every subject may be stretching the capacity of schools to educate students with sufficient depth in critical fields like language arts, science, and mathematics.

We recommend that the standards review process include an examination of the need for standards, curriculum frameworks, and statewide paper and pencil assessments in all of the content areas. We believe that all students need to have a well-rounded education to prepare them for successful adult lives. But within that broad education opportunity, priority may need to be given to some subjects because of their increasing importance to all adults.

However, one cannot address the State's assessment system without looking beyond that to other factors that impact student achievement. And the areas we're concerned with include the standards, teacher professional development, and accountability.

Under the topic of standards, we've already heard that Achieve recommends that the standards be made clearer and more specific.

I'll skip down to teacher quality, because we've heard so much about the Achieve report already.

Few would dispute that the quality of the teacher is an important factor in student achievement. A growing body of research shows that teacher quality is more critical than even formerly assumed. To ensure that every student achieves high standards, teachers need to be prepared for the challenges that are generated when the educational bar is raised.

The State needs to develop and disseminate data about the quality of New Jersey's corps of teachers. Widely referenced studies indicate that a deep understanding of content matter is critical to good teaching, yet many school districts allow individuals to teach in areas that are not within their specialty area, particularly in the field of science and math. While limited data is

available, some estimates suggest that as many as 37 percent of New Jersey's teachers are teaching out of their field, as defined by their college major.

The educational system needs to increase its investment in teacher professional development, raise the quality of knowledge and skills offered in professional development, the amount of time provided to teachers for skill development, and ensure that professional development content and skills are strictly aligned with State academic standards. The State needs a high-quality and stable professional development system. Professional development programs should improve the ability of a teacher to help all students reach high academic standards. Professional development should be expanded to include mandates and service delivery systems for school personnel such as superintendents, principals, and instructional supervisors, as well as school leadership or board officials.

Specifically, the State needs to establish a system of professional development centers, local hubs where teachers and school administrators can obtain the best professional development possible, centers where educators could go to obtain the best thinking about professional development and classroom practice and receive direct professional development service. These centers would deliver much more than the one-week summer workshop as professional development experiences, but would extend service to include ongoing observation, support, and follow-up with the teacher in the classroom, assuring that the best research-based classroom practice is available to the students in the system, which is our future workforce.

It is time to recognize that this admittedly expensive system might be the best investment the State can make in assuring that all students achieve

the core curriculum content standards. California is currently establishing such a system and investing close to \$40 million in the effort. Soon, any California teacher will be able to receive top-notch, in-depth, State-supported professional development, complete with classroom support and follow-up. Our current system of hundreds of providers delivering locally selected teacher training in short and unconnected increments will not allow us to keep up. Tangentially, such a system would also provide New Jersey with a great teacher recruiting and retention benefit.

And finally, on the subject of accountability, investments in costly assessment systems are wasteful if parallel accountability systems and data tracking systems are not in place. Accountability systems that do not include continuous improvement mechanisms become punitive rather than positive exercises. Accountability systems must include timely data dissemination in formats that make sense to the public, the parents, and educators.

Accountability systems must also be fair, that is, based on relevant comparisons. Student performance data should be disaggregated for mobility factors so that education stakeholders can evaluate the performance of students who have been in the district for a significant period of time against the performance of students who have been within the district for only a short period of time. Data disaggregated in this fashion will also permit school districts to respond to the needs of transient students most typically found in urban areas, where mobility rates are the highest.

Longitudinal tracking of student performance data can only be accomplished through a statewide system of unique student identification numbers. New Jersey does not have such a system, but many states do.

Therefore, comparative analysis among schools that can define best practice, reflective self-assessment within districts that can drive curriculum selection and teacher training, and measurement of improvement over time become difficult statistical problems. The Department of Education should implement a student identification system as quickly as possible.

Finally, we believe that high achieving organizations motivate performance through a system of incentives and rewards. We believe that disinvesting in poorly performing schools to encourage improvement is questionable practice. While it may be more difficult, finding ways to involve all education stakeholders in systemic reform ensures that all students, including those without educationally savvy parents, will have the opportunity to learn.

I am pleased to provide you with this summary of what is -- of the document that is attached, *Getting Down to the Business of Learning*, and would be happy to answer any questions that you might have.

SENATOR MARTIN: I take it you would consider, perhaps, a surcharge on the corporate business tax in order to help fund that expensive proposal -- the \$40 million, California-like program. (laughter)

MS. EGRECKZY: No, we often hear that, but I suggest that we need to look at this as a system.

Right now, there are millions -- literally millions of people who are relying on corporate performance to sustain them in their retirement. They're all planning on 401(k)s to sustain them. They're all expecting corporations to provide dividends to stockholders.

Secondly, corporations, as they are now, pay significant amounts of taxes.

And finally, as you just heard, Prudential itself gives \$25 million away in its philanthropic efforts.

So corporations are already providing significant levels of support -- direct financial support to everything that's going on in the State.

However, what we're suggesting is that what we need to start doing is prioritizing the way we're spending some of our money.

SENATOR MARTIN: Well, we're always prioritizing. The question is reprioritizing.

MS. EGRECKZY: Reprioritizing.

SENATOR MARTIN: So I take it the answer is no, we can't look to the corporate community to come up with any more funding for those proposals.

MS. O'MALLEY: If I may, the Aptech Training Centers that were established in the state about four years ago were established with corporate help. And we did this not through direct funding of these centers, but providing professional assistance to the centers in helping with setting them up and helping getting them appropriate hardware and software and then providing training.

So perhaps rather than giving direct money through tax assistance, we can provide this kind of professional and in-kind assistance.

SENATOR MARTIN: Part of the problem is developing programs that work. We've had SSI in the State for several years, which was directly supposed to promote teacher professional development in math and I think also science, but certainly in math. It's coming to a close, and under the grants that were originally provided, I'm not so sure that that program delivered all that it was set out to do. And it may be an example. It's one thing to say we'll just set

up this great center and people will go there and then they'll all be professionally enhanced and be able to solve a problem. It's more complicated than that, as most educational issues are.

Lord knows, New Jersey has spent, and this I'm sure the corporate community would agree with-- We have spent a lot of money in the State. And I think most of us feel we probably should do a better job. All of the state should share responsibility for that. But we haven't got as much value for our money, I think, as we would have hoped.

MS. EGRECKZY: I agree.

SENATOR MARTIN: How we do a better job, of course, is the multibillion dollar question.

MS. EGRECKZY: But, Senator, you should also know, for example, that occasionally I am asked, as a Chamber representative, to deliver a speech at a teacher gathering or at any function. And I go in. I give a 15 or 20 minute presentation, and teachers are getting professional development credit hours for listening to me.

Now, I find that appalling, because there's nothing I'm going to say to a teacher in 15 or 20 minutes that's going to change their classroom practice, so I think we really need to look very much--

SENATOR MARTIN: You should give them a test at the end.
(laughter)

Your point is well-taken. Our whole-- The notion of professional development really needs work in this state. And a lot needs to be done, and we need some better recommendations. And then we have to act on them.

Thank you, both.

Thank you.

I'd like to hear from some representatives of the New Jersey Association of School Administrators.

Dr. (*sic*) Groff, Trudy Doyle, Robert Bumpus. Are they still here?

MR. CANNON: Yes.

SENATOR MARTIN: I'm trying to--

You've got a new role there, Doug.

DOUGLAS GROFF: Yes, I do.

Is it working? (referring to PA microphone)

SENATOR MARTIN: It's kind of like keeping up with players in the National League. They sort of move around.

MR. GROFF: I was traded. (laughter)

ROBERT L. BUMPUS: He's a free agent.

MR. GROFF: I was a free agent.

First of all, we'd like to say thank you very much, Senator, and to the members of the Education Committee.

I was a former assistant commissioner, so it was nice to hear the new Commissioner here today and his thoughts.

Currently, I'm the Superintendent of Schools in Galloway Township in Atlantic County. So I did say hi to Senator Gormley, who recently visited our brand-new middle school. So it's an exciting opportunity for me, with my two colleagues here, to offer some brief testimony to the Committee.

So we say good afternoon.

And I'd like to introduce Bob Bumpus, who is Chair of the Curriculum and Instruction Committee for NJASA.

MR. BUMPUS: Good afternoon. My name is Bob Bumpus. I'm Superintendent of the Cumberland Regional School District, which is a regional high school in Cumberland County serving seven constituent districts. I'm also Chairman of the Curriculum and Instruction Committee of the NJASA, which is the primary professional association for high-level administrators in New Jersey.

And I have a colleague with me who has worked very closely with Doug and I on a number of committees, and she is Trudy Doyle.

TRUDY DOYLE: Nice to meet you all.

MR. GROFF: If I could, Senator--

What we're going to do, and we're going to do this briefly, because we know that the hour is late, and it's been rigorous testimony from a variety of stakeholders--

The first thing that we want to do very quickly is review the New Jersey Statewide Assessment Advisory Panel's final report. Last May, Commissioner Hespe reached out and asked if I would chair a large group of about 27 -- 29 people from across the state to take a look at how we were going to assess world languages, the visual and performing arts, and also the physical and health requirements of the core curriculum content standards, because, as the Committee knows, they are the next areas that will have to be assessed as part of the core curriculum content standards, and also the timetable that's been set by the Department of Education.

I believe, at that time, Commissioner Hespe realized that we had four that were out there -- three, plus a field test for social studies -- and that there was large concerns from the field that adding on three more State tests,

from the State's perspective, in any of these grade levels, fourth, eighth, and of course, eleventh grade, was going to add to the rigors already that we were going through.

The Commissioner did meet with the task force and opened it up and said to think outside the box and try to come up with some new ideas on how we could best do this.

Without a doubt, this group met 13 times from May to December, and we submitted a report to the Commissioner. And I believe that report was shared by the Commissioner with all the major stakeholders -- educational stakeholders, who had an opportunity to make a comment on that. And if the Committee does not have it, I'm sure that that could be made available, because it was made public.

So, at this time, very quickly -- and it's the one page of recommendations-- I will say this was a very, to say the least -- and I would ask Trudy and Bob-- This was a very contentious group, because we were made up of superintendents, principals, teachers, experts in each of these areas who felt very, very strongly about their particular area of expertise. Dana was on there representing the business community. We had a professor in testing and assessment from Rutgers University. So it was a very broad group, and we were able to come up with a report. And we called it a consensus report. And these are the major recommendations.

The panel recommended that the three assessments share the following characteristics: That performance assessments should be developed collaboratively by the State and districts and administered and scored locally under uniform standard conditions. Initially, performance on these three high

school assessments will not prevent a student from graduation. However, course requirements must be met.

The assessment should provide for ease of administration and reporting.

SENATOR MARTIN: Just so we understand, that's a big change.

MR. GROFF: Yes, it is.

SENATOR MARTIN: You're taking the high stakes off. That would be your recommendation.

MS. DOYLE: In those three content areas.

MR. GROFF: In those three content areas, which would be visual and performing arts, health and phys ed, and world languages.

The assessments should measure what students know and are able to do as defined by the core curriculum content standards. The assessments should be developmentally appropriate in both content and format. Public reporting of the results should be based on common performance standards. The assessments should be supported by content-specific professional development that addresses both instruction and assessment.

This is a very important one, Senator.

The assessment should be integrated into the instructional, ongoing learning program. And the Department should implement a district and school accountability system to annual monitor results. And I'd like--

SENATOR MARTIN: Integrated meaning you don't want--

MR. GROFF: Pardon me?

SENATOR MARTIN: These would not be special -- separate courses. What does integration mean?

MR. GROFF: Integration is-- What we're trying to do is come up with a commonsense way of doing assessment in the State of New Jersey. I think everybody wants that. With these three areas, I think we have that opportunity -- where you could actually do it within the ongoing instructional program. In other words, if these -- if the standards were developed, and the test specifications were developed jointly by the Department and people in the field, they could be given during the regular instruction process. It wouldn't have to be, necessarily, in fourth grade. It could be in fourth, it could be fifth, it could be third grade, because we do put a lot of pressure on fourth grade right now.

The idea is that these are-- We assess students every single day. And teachers know best when to do that. And we have something in place right now called the QAAR, which every district must submit to the county superintendent, which is then reviewed, of course, by the Commissioner. We could have a natural monitoring -- another two -- let's say two pages that could be added onto this QAAR, which is already there, which would then allow each district to verify that they have given the assessments. And it would be signed off by the teacher, the principal, the superintendent, and the county superintendent of schools. So we already have a monitoring item in the field that's used constantly now to rate the districts.

SENATOR MARTIN: You heard Commissioner Gagliardi in response to my question about whether he was going forward with additional tests which would increase the time. And he did not give a definitive answer. I take it it's because of recommendations like these where he may be trying to consider alternative ways of meeting the criteria, which is why he's sort of

maybe not reached his mind. Wouldn't you sense that that's sort of, at least, part of his thought process?

MR. GROFF: Definitely, without a doubt. I mean, Commissioner Gagliardi has only been on the job a short amount of time, and it's a huge nut to crack. And if you take assessment, I know that there's a lot of research and a lot of groups that are meeting. And I know that in looking at this report, this would probably help him or be a part of what he would look at as possible recommendations, because I know he knows that they do have to be assessed. And he is trying to really take a look at how is the best way to assess them without placing burdens on districts and on students and on teachers.

MR. BUMPUS: Please keep in mind, as high-level administrators, we are charged with implementing whatever the State requires us to implement. We would have to do it systemically.

Our major concern is that this thing is getting so broad and so cumbersome that we're not going to be able to assess well, particularly the validity of the assessment. And these areas, in particular the integration part in the high school, may work like this.

Right now, there's a 10-credit requirement in world languages for every incoming freshman next year. And within the context of the course -- maybe the first-level course, certainly the second-level course -- the assessment would be provided, and the students would meet the requirements, via the assessment, but it would be within the context of the course. In our school in particular, it's a box scheduled school. That student could manage that assessment within his first year of high school, because it's on a semester basis. So he wouldn't have to wait until his junior year. So it would really narrow the

high-stakes assessment that he would have to pass in order to get a high school diploma.

MS. DOYLE: And just -- quick example for the Committee on the performance assessments we're recommending for those three areas -- might look like world language, where two students would get up in front of the class and create a dialogue and speak with one another. And they would do that as part of their regular classroom day. And the teacher would use her rubrics to score that dialog. And that would be the assessment for that piece. That's just an example very different from what we think of in terms of paper and pencil tests.

So, for those three areas, this committee made that recommendation to the Commissioner -- that we look at authentic and performance assessments in those three areas and to keep them in the context of the instructional day. And it would, obviously, be something done by local districts. And we're going to address some of that in the next part.

MR. GROFF: If I could just -- with the additional recommendations -- because this was a very large undertaking. As I said, we did meet from May to December, and we met 13 times. And people came from across the state. And we had very good attendance. So this was something very near and dear to everyone's heart. And they didn't really miss many meetings.

The committee of the whole agreed upon the following additional recommendations: The Department should study programmatic impact data in these three content areas as a result of these proposed assessment changes; the Department should reestablish Academy North, Academy Central, and Academy South, as much as the business representatives talked about a center -- a learning center -- to facilitate needed professional development.

When I was at the Department, we were very positive in trying to get the academies back. I think the best compromise was the new virtual academy. And I think it's very good. But we have to remember that all teachers don't have access to all computers all the time, so I think it's a step in the right direction, but we need to broaden that a little bit more.

The Department should convene a task force to research and study the uses and applications of technology in assessment, such as CD-ROM portfolios. I think, Senator, you brought that up earlier.

The Department should continue its leadership role to facilitate the implementation of the assessment panel's recommendations in these three content areas.

The Department should recognize that these three content areas are different and distinct and that they should be addressed individually because of their unique characteristics.

The assessment system developed for these three content areas must include students with disabilities and limited English proficient students pursuant to N.J.A.C. 6A:8.

The Department should consider benchmarking the standards at smaller grade clusters, for example, Grades 1-2, Grades 3-4. This can be accomplished during the standards review process beginning in May 2001. And I think that's -- it's starting right now. Commissioner Gagliardi has already indicated that he has commissioned a committee to start looking at those.

So we feel that this was a wide-ranging group of professionals representing all the stakeholders in education. And we were pleased to submit these recommendations to Commissioner Hespe at that time and offer these, of

course, to Commissioner Gagliardi, should they help him and the Department meet this rigorous challenge in assessment.

Thank you very much, Senator.

And I turn it over now to Bob Bumpus.

MR. BUMPUS: I might add that Commissioner Hesper was very much in favor of our -- of this final report.

Also, I'm referring to this document that I've handed out earlier. And we have further concerns and issues we'd like to share with the Committee this afternoon.

The first one is, as Chair of the Curriculum and Instruction Committee, from time to time we have recommendations that we give to the executive committee. And our latest recommendation was an incentive for all school districts. And it would permit an exemption or a waiver of the administration of the ESPA if a school district was compliant or proficient on the GEPA in all three areas, or all the tested areas, for three consecutive years. That would provide an incentive for school districts to do a good job. And I was interested in Commissioner Gagliardi's comments about the high-stakes nature of our testing. And he mentioned that the high school proficiency assessment is the highest stakes assessment. The other two, the ESPA and the GEPA, are certainly required, but the high-stakes issue -- we really need to address that.

Other concerns and issues: State testing should be limited to language arts literacy, mathematics, and science, as far as the traditional paper and pencil test; the ESPA still needs to be reexamined for alignment with proficiencies and to ensure that it is developmentally appropriate for fourth-

graders; districts should be permitted to use nationally recognized tests in Grade 4 in lieu of the ESPA.

And the major difference between a national standardized test is that it's a norm reference test. It's based on a normal curve. And it's really-- It really distinguishes the different levels of academic achievement, as opposed to a standards-based test, which is based on standards, which requires everybody to eventually hit the mark. So there's a big difference between the nature of those two tests.

Proficiency levels on ESPA, GEPA, and HSPA should not become part of the State Department of Education monitoring process until districts are given sufficient time to provide teaching staff with appropriate training and field testing to provide greater validation.

Supplemental to the absolute monitoring standards in fourth grade, which is 75 percent -- and GEPA is 75 percent, and HSPA, 85 percent -- should be progress indicators, which allow school districts to meet those marks over time. I know that the Department has something called yearly annual progress, which is based on a seven-year cycle -- 3 percent growth over seven years. So that should be a consideration, also.

The State Department of Education should reestablish the academies, as was mentioned in our report that Mr. Groff gave.

The State Department of Education should lead and fund the development of performance assessments. That was also mentioned. Local districts should then be given the opportunity to select the performance assessments from the State-developed menu. The performance assessments could be used for all core content areas, but most importantly should be used

for the four remaining content areas: social studies, visual and performing arts, world language, and physical education and health. I might add, practical art should be included with the performing arts, also. Passing a set level of proficiency in the four content areas should not be a graduation requirement at this time.

I defer to Ms. Doyle's--

MS. DOYLE: Student performance in the four remaining content areas should be reported locally and monitored, as Doug had mentioned earlier, through the existing QAAR process. That is a system that is already in place, and this would just be adding a piece to it. This is an instrument -- an idea that came up during this -- the Commissioner's Statewide Assessment Committee, and we have a lot of input from a gentleman from Rutgers University on the possibility of using matrix sampling, which could be used as a viable alternative to ESPA or GEPA to audit program implementation rather than individual student performance.

The reason we talked a great deal about matrix sampling in our Committee was because it gave us an opportunity to shorten the test. However, in matrix sampling, you do not get enough information or data to look at individual student performance. However, if it's the State's goal to make sure that the core curriculum content standards are being implemented, matrix sampling would ensure that. So that's another thing that we had talked about -- hadn't really gone anywhere at this point.

MR. GROFF: Trudy, thanks, because I know where this is going to lead.

That was one of the largest discussions that we had. In other words, in fourth grade, do we really want to take a look at individual students, or K-4 -- should that be setting the curriculum and the patterns and really taking a look at teacher proficiencies and setting a curriculum that then kicks in -- of course kindergarten to four -- but where you could have a test that really sort of takes a look at the curriculum that you're using. And then you would start more of the heavier or more high-stakes testing starting in eighth grade and moving into the high school using-- In other words, coming along with something using matrix sampling. In other words, that's what NAEP does. NAEP takes a look at, really, building level and how is your building doing, how is your curriculum put together to meet the various standards. So that was something that we talked about. There wasn't a clear consensus, but it was something that we were very interested in as maybe taking the high pressure off and looking at curriculum development and aligning their curriculum to take a look at that at the fourth grade.

MS. DOYLE: We also would encourage the Department to pursue the use of technology -- we mentioned this earlier -- to provide students with the opportunity to self-assess at regular intervals. We currently have the ability, using technology, to do this.

I'm sure many of you are aware of the fact that you can now take the graduate record exam and the SAT on a computer and get your results almost instantly.

Providing students with immediate feedback on their mastery of proficiencies would really be a great step in the right direction for us. However,

it's something that would need-- Obviously, it would have to be a priority, and it would need to be developed at the State level.

SENATOR MARTIN: But you couldn't do it for open-ended questions.

MS. DOYLE: Yes, you could. Well, I don't know. Maybe you still could. As long as there was a rubric in place to measure the open-ended question response, I think you might be able to-- It would be a little bit trickier, but a lot of computer software companies are developing software that has built-in intelligence.

What we're saying is that this is the moment in time--

SENATOR MARTIN: You're right, you could do it.

MS. DOYLE: --to explore it. But that's a good point.

And again, we would emphasize that we believe, as educational leaders, that assessment should be, and can be, integrated into regular instruction.

The core curriculum content standards, as you heard many times today -- and you know this initiative is now taking place at the State level -- need to be refined and integrated. The indicators should be reduced in number to reflect the primary concept and skills of each discipline.

And let me just speak to integration for one moment, because this was important in our committee. During the committee, we talked about the fact that the tests were so long. We talked about the fact that it could be very appropriate to have a reading passage on the test that was based in the social studies discipline. And at that time, we were told by the Department that that would not create valid data, because we were mixing two content areas.

Did I say that correctly? (affirmative response)

As educators, we are trying to teach in an integrated fashion so that children see that social studies is not a separate subject from life, mathematics is not a separate subject from life, and therefore, some of our testing could be done in a much more integrated fashion. And so, again, we would encourage that as an initiative for the future. If we're encouraging teachers to teach in an integrated fashion, we believe we should also test in an integrated fashion.

MR. BUMPUS: By the way, brain-based research into learning -- it indicates that's the way we learn. That's the way the brain operates. The brain operates in a connected, integrated way. So we just want to get in line with what we're learning about the brain -- physiology of the brain.

MS. DOYLE: And we've also asked that there be a professional oversight committee formed with New Jersey Department of Ed members and field practitioners to provide a continuing monitoring of the assessment program.

And as you also heard today, for the educators to track students and track improvement in student achievement, which is our goal, we do need disaggregated test data for special education, limited English proficiency, and mobility rate and exempting students from benchmarks used for monitoring and publishing those test results.

The mobility rate, I know you heard a great deal about. It is very important that districts be able to really look at their programs. And in order to do that and improve their programs, they need to look at how many of the students that are being tested have been in their program for a minimum of two years, at which time you could maybe -- hopefully make a difference.

Any questions? We went very quickly because it's--

SENATOR MARTIN: It was a lot of good food for thought.

MS. DOYLE: Well, we'd love to come back if you ever want to invite us. I'm sure that the three of us would love to have a little more time with you.

MR. GROFF: Thank you, Senator.

SENATOR MARTIN: Thank you.

MS. DOYLE: Thank you very much.

SENATOR MARTIN: I think we're going to hear from two more people. These are individuals who I don't believe are here on a regular basis. The other representatives from school boards and NJEA and NJPSA and Garden State Coalition, we will reserve for a later time, but the hour is growing somewhat late, and our attention can only absorb so much at one time. But we did want to hear from Wendy Webster Odell, (phonetic spelling) if she's still here. I'm not sure-- Is she here?

UNIDENTIFIED SPEAKER FROM AUDIENCE: I don't think so.

And if not, Len Mitnaul-- Is he still here?

Sir, thank you for coming.

LEN MITNAUL: First, I would just like to thank you for your interest in education and the children of New Jersey. I applaud you for valuing education and that you're trying to improve the schools in New Jersey.

My name is Len Mitnaul, and I've been a public school educator for over 20 years, and I've taught in both suburban, urban, and rural areas.

My chief focus is early childhood education, specifically kindergarten and first grade. I've also taught the dreaded fourth grade. I have degrees in early childhood education, elementary education, music, and mathematics. I'm currently working on my doctorate degree at Teachers College at Columbia University.

SENATOR MARTIN: So am I.

MR. MITNAUL: Originally, I had planned on presenting a position paper -- you know, scholarly type work that Columbia would be proud of -- with many pages and references and so on. But I know how much paper you already have. So instead, I decided to just present it to you in Mother Goose language. That's what I tell my kids.

First of all, teachers have always used a variety of assessments and strategies to help them evaluate student progress, improve student progress, and to make sound instructional decisions.

A good assessment system includes three essential elements. -- different types of informal and formal assessments. No. 2, it should help the teacher integrate and improve instruction. And No. 3, both the teacher and the student should self-assess. That should occur throughout the learning process. I find that the ESPA does not do that.

For example, one of the things that standardized tests ignore is the many characteristics of good learners and excellent students and productive citizens. For example, they don't measure initiative, creativity, conceptual thinking, imagination, effort, respect, judgment, commitment, good will, caring, self-worth, ethical reflection, self-control, understanding, and just so many things that are not actually on the ESPA.

What they can measure and count are isolated skills, specific facts, and specific functions. These are the least interesting and least significant aspects of learning.

In order to understand what students have learned, we need to look at what they have learned in different ways throughout the year. High-stakes testing can be detrimental to learning and undermine many practices that have been so critical to successful teaching practices.

I think if we can give teachers a timing within the framework of the State's standards, our chances of creating schools that are meaningful places for children to go each day will be increased.

The National Research Council described the negative effects of achievement testing on instruction. Tests become an end, not a means to assess educational objectives. Knowing this, as you've heard before, teachers often teach to the test, not to the curriculum or the children.

I'd like to give you an example. One of the things that standardized tests do is actually punish the thinking test taker to the point that some teachers actually advise their students to, in effect, dumb themselves down so they can do better on the test. One science teacher explained to me how it's necessary to unteach some things when test time rolls around. If we teach creative thinking, if we teach analytical thinking, and if we teach children to look beyond the obvious into the not quite so obvious, then a proficiency test comes, and they think analytically. They think beyond the obvious. And they will get the questions wrong, because they know of all the exceptions to the rules. And they say, "But what if--" And so I try to reserve at least a week prior to the

proficiency exam to throw everything aside. And they check their books. And we close down school. And we give them proficiency drills.

Another teacher does much the same thing, telling students, "You just have to think, what's the least dumb answer here? Do not try to use your mind on those."

One of the things by publishing test scores-- It basically just tells us the socioeconomic status and the available resources of towns and communities. Testing does not tell us anything about what happens in classrooms or the quality of teaching.

I would ask each of you on the Committee to actually take the ESPA for yourselves, and the GEPA, and to have your test scores published as schools do.

I'd like to answer some of the questions -- not questions, but some of the things that were brought up and I had written some notes about. And I just want to respond to them.

One of the things that we do is, we keep looking at the product -- the students, but we don't really look at who's providing or who's producing the product -- in other words, the teachers who keep saying test scores are low, we need to increase this. But we really need to look at what we can do to improve teachers.

Another comment I'd like to make is, I don't think we should be proud that we are one of nine states that look at high-stakes tests for promotion. I think we should be ashamed of that. It's very sad when schools can spend 13 or more years educating students for it to be boiled down to one test for graduation.

Just imagine that in order to keep your job or your position, it was based on one test.

SENATOR MARTIN: Don't you have to pass a certification test in order to get your doctorate?

MR. MITNAUL: Yes, but--

SENATOR MARTIN: It's one test, right? You don't pass it, you won't graduate.

MR. MITNAUL: I agree, yes. But I don't think we should ask that for our children.

SENATOR MARTIN: Just say that for -- so when they go all the way through graduate school and try to get a doctorate and they fail, that's when we should lay it on them?

It's just a thought.

Go ahead.

MR. MITNAUL: Any other questions or comments?

SENATOR TURNER: Where do you teach?

MR. MITNAUL: I teach in Trenton.

SENATOR TURNER: Which school?

MR. MITNAUL: Pardon?

SENATOR TURNER: Which school?

MR. MITNAUL: Woodrow Wilson. I teach first grade.

And I know Trenton scores do not look well, but I just want to bring up one point that--

I have another colleague in first grade. And we enter the reading rainbow contest for writing stories. And it's a national competition for writing.

And I don't think it's published anywhere that -- the first runner-up -- there's only two winners -- came from our school. But if you look at our ESPA scores, you would say the students aren't writing.

SENATOR MARTIN: You referred to a list of areas which, I think, many of us -- at least I would agree -- are areas that-- They should be, if not taught, at least sort of encouraged and valued in schools. I think the difficulty we face is trying to measure how you would go about determining whether young people have achieved those goals.

Are you aware of some ways in which assessments could be done in those areas other than sort of a teacher's sort of general observation?

MR. MITNAUL: I think one thing we need to do is to not have just one high-stakes test. In other words, we look at students -- how they're doing in September -- look at their writing, how are they speaking, what talents they have -- and then to assess them again in the middle of the year, and then again at the end of the year, so we can actually see, wow, these students have made progress.

Someone made the comment about -- it's really not fair that you have schools that are -- you can say are privileged schools, and the students are perceived to be doing well, but they're really not making a large amount of progress, and you might have schools in areas like Trenton where it seems like the scores are very low, but the students have made lots of progress. And that's not really seen by the public. They just look at the test scores, and they go, "Oh, Trenton's pretty bad." And that's what they see.

SENATOR TURNER: Well, maybe they should be tested when they first start school. And then when they take the fourth-grade test, you get

a better picture, as far as where have they come from the time they entered the school system.

MR. MITNAUL: I think one of the things that has happened with testing is-- For example, we have a lot of students that transfer in and out. And when you have a student that transfers in and it's close to test time, a lot of teachers are going, "Wow, this is really a liability. This student doesn't know these skills -- into the student quickly."

I think in any other profession-- If you were a doctor, and as a patient you came and you said, "Well, I have cancer," and if the doctor's records were published -- how many patients died and how many survived -- I think the doctor would have a different outlook on, "Well, should I really take this patient, because it's going to make my scores look bad?"

SENATOR MARTIN: You know, that's a good example. It is difficult to make measurements across school districts. And that's a very good one. I've also, at least, been advised by some medical professionals, that's exactly the way some hospitals discourage people, for example, with heart bypass operations and others which they now tell by certain statistical data, "Our hospital system has a 95 percent success rate." And in some cases, some of those hospitals, at least allegedly, discourage people who seem to be at a higher risk to either use that hospital or to take certain operations, because they also have become obsessed with the results. And so the results drive the medical practice, which is not something most of us relish, and the way in which they perform.

But it is perfectly understandable, I think, that a school teacher in Trenton or a school teacher in Morris Plains or a school teacher in Bergen

County, knowing that it's going to be a reflection on them, even though some of these areas are outside of their ability to influence, that are going to try to take some actions that would maybe -- I don't want to say artificially-- But they're defensive mechanisms. And that's pretty understandable, given how much pressure may be brought upon them.

SENATOR BAER: Senator, I made the very same observation that you did, in terms of the analogy with hospitals, which I think use sometimes risk-adjusted statistics to take into account the differences in the degree of challenge or difficulty that one hospital might have with its cases as opposed to another.

SENATOR MARTIN: I think you have a lot to offer with some of your recommendations. I don't know that we know quite how to absorb them at this time, but I would appreciate if you would give us some more writing that we could look at and, at least, reflect. And I'm not sure how, but I think if some way you could contribute to the Department -- and if there were some way you could engage-- I think some of your comments are, at least, worthy of attention.

I really appreciate the fact that you would come here and testify at our hearing.

Anyone else?

SENATOR TURNER: Absolutely.

Trenton is very lucky to have you as a teacher.

SENATOR BAER: I just wanted to say one other thing. In case it seemed to you that I missed part of your testimony, I'm going to catch every word on the computer audio feed this evening.

MR. MITNAUL: Thank you.

SENATOR BAER: Thank you.

SENATOR MARTIN: Thank you.

That will conclude our--

And I appreciate the others who were going to testify. We can catch you at our next Committee meeting -- give it a little more attention.

(HEARING CONCLUDED)