

**COMMONWEALTH OF MASSACHUSETTS  
BOARD OF EDUCATION  
\*\*\*SPECIAL MEETING\*\*\*  
CROWNE PLAZA HOTEL  
ONE WEST STREET  
PITTSFIELD, MASSACHUSETTS**

**MONDAY, APRIL 24, 2000  
4:00 p.m. - 6:00 p.m.**

**MEMBERS OF THE BOARD  
OF EDUCATION PRESENT:**

Mr. James A. Peyser, Chairman, Boston  
Dr. Roberta R. Schaefer, Vice-chairperson, Worcester  
Ms. Patricia Crutchfield, Southwick  
Mr. William K. Irwin, Wilmington  
Dr. Judith Gill, Boston  
Mr. Marcel LaFlamme, Chair, Student Advisory Council, Monson  
Dr. Abigail Thernstrom, Lexington

Dr. David P. Driscoll, Commissioner of Education

**MEMBERS OF THE BOARD  
OF EDUCATION ABSENT:**

Mr. Charles D. Baker, Swampscott  
Dr. Edwin J. Delattre, Boston

**ALSO PRESENT:**

Nancy Catuogno Varallo, Registered Diplomate Reporter

**CHAIRMAN PEYSER:** Welcome, everyone, to Pittsfield. This is, as you know, a discussion meeting, it's not a business meeting. There will be no votes, no decisions taken. Hopefully we'll have an informative presentation and discussion around school and district accountability and, without further ado, let me turn it over to the Commissioner who will then hand it off to Julianne Dow.

**COMMISSIONER DRISCOLL:** Mr. Chairman, in the interests of time, I will turn it over to Julianne. School and district accountability may very well be the most important driver for education reform. That's been the experience in other states. Certainly the emphasis on MCAS, particularly here as we go towards the high stakes, gets all the headlines. But it is a good, sound school and district accountability system that I believe really does mobilize the people in schools. So with that, I will turn it over to Julianne.

**MS. DOW:** Last September the Board of Education adopted a framework for the school and district accountability system. We spent more than a year before that looking at different approaches to school and school district accountability from around the country and other nations. We have looked at our own frameworks and our own testing system to formulate a plan that would work for Massachusetts. Since September, we've been working on developing components of the school accountability system, beginning to prepare the criteria and standards for a district performance evaluation, and doing some initial piloting of the standards and protocols that we have in mind for evaluating our districts.

I'd like to begin with where we are in the district performance evaluation process and look at the proposed standards for district performance evaluation that we have brought for this Board meeting. We think we are ready to go out for a broader public engagement or discussion about these standards. And if the Board were in agreement with that, that would be the next step.

**CHAIRMAN PEYSER:** Should we let you go through the district portion and then have questions and discussion, or would you rather people interrupt you as you go through it?

**MS. DOW:** I'm going to present very little. I would like us to spend the time talking. I want to set the larger framework and I welcome any interruptions on that. I think we should talk about the district and then separately talk about the school process and identify where the break is.

**DR. THERNSTROM:** If it's going to be brief, why don't we let her go through the discussion of the district and then ask our questions?

**MS. DOW:** It's not so much a question of the things we need to look at in evaluating district performance, but rather how to group and categorize them so we can effectively analyze them and so that people can understand what we are doing.

In January we circulated an earlier version of standards, and there were 76 standards. We used that document and the draft protocol that we had to begin piloting in seven districts this spring. There are three teams this week in three different districts working from that earlier version. One of the things that the teams came back with is that there are too many little pieces. We need to look at this through some larger views and we need to be able to then draw up our conclusions and render our judgment on those larger chunks of the information about district performance. The version that's in your Board book represents this new level.

What has been the most challenging about this is that the operation of a school district is very complex. To deliver quality education in the schools is an immense and complex undertaking. What you're looking at really describes six major categories of inquiry that we would look at. And here we've put beside them the relative weight of those sections of inquiry. So it is very clear that when we look at Curriculum & Instruction in the Post-secondary Preparation for Work and Employment, that's the area in which we have the greatest weight. That's where we are looking at the student results that are obtained in each of the subject areas. It's where we are looking at the graduation rates and the dropout rate. It's also where we are looking at what is being delivered in the schools and the district in order to achieve those results or to support those results by students. And that has the largest weight and therefore will also be the area where we devote the most time and energy in the review process. The next three boxes, Teacher Quality, Learning Environment, and Student Support Services also have to be analyzed at a school-by-school level as does the Curriculum & Instruction. So all of those five have to do with what's going on in the schools in the district.

The bottom green box, Organizational Effectiveness and Efficiency, is where we look at how the district operates as an organization to help deliver a high-quality education and support the attainment of high student results. So this is the top-level view. Underneath these "super standards" are sets of Indicators. Then we have the identification of specific evidence that we'll be looking for, sources of evidence, places where we'll look to find evidence; and then measurement criteria for how we'll consider that evidence and put it into performance levels or categories similar to what we've done in looking at the range of performance of students. As you start to look underneath for these Indicators and evidence and sources of evidence, you can pull away those larger headings, and this is what you see at the next layer down. We have these categories of Indicators. And within each of those are some sub-statements. And in the document that you have, you have these Indicators. So this is just to get a sense that although we have been able to express these, I think, in a more understandable form by creating those larger groups, underneath it is a lot of substance. I don't want anyone to think because we have now expressed it in terms of these six larger categories, that the substance has gone away. We have been working very hard to develop the detail underneath. And we have completed the most complex task, something that the teams been troubling with, that is, the Curriculum & Instruction area.

In some districts they may have a unified curriculum across all their schools, and in others they may have curriculum that's different in all the schools, although it all may be aligned to state standards. So we have to look at schools. But within schools there are special ed. programs and bilingual ed. programs, and regular ed. programs, and they don't necessarily have all the same curriculum, they don't necessarily have the same instructional practices. So we have to look by program. We are looking by subject area and then we are looking at Curriculum, Instructional Time, Instructional Practices, and Student Performance Assessment. There are several dimensions. We've concluded that we will need to ask every district to review the quality of its own schools and compile the evidence they would then present to us. Then we'd be in the situation where we could spot-check and verify. But to start from scratch and go out on site and do this analysis is a huge undertaking. By developing a protocol for what it would take to do that we are saying to the districts, "This is a starting point for you to use, if you wish, as a way of looking at your schools and gathering the information we will need for a district review." As I mentioned in that earlier version of the document, we had used as an organizing principle these three categories: student results, quality of educational programs and services, and organizational effectiveness. We'll be looking at those as we look

at the evidence. We're going to look at those six items we had on that first slide, but each time we'll be looking for student results. Or in some cases it's other results. For example, hiring practices, it's actually results in terms of who your staff is, who you have been able to hire and retain. So we'll be looking not only at the processes, but at the outcomes of processes as sources of evidence.

Again, we'll be looking at different kinds of data. We'll be collecting some data that we have not historically collected around teachers and educational personnel in the schools, but also we are going to be gathering a variety of documents that we have not historically looked at and studied at the district level. And then a portion will be actual on-site interviews and observation. But a lot of work will be done in this process before a team goes out on site looking at the data and documentation. I just have given here one example so you can see the layers. I don't use this example because it's perfectly developed or that I'm confident that we have all the right Indicators and all the right evidence yet. That's very much what we are in the process of wrangling with now.

In the area of Teacher Quality, which is, again, one of the standards that you have in the document, under Key Question there are four Indicators, and within each of those Indicators it goes down another level. Let's look at one--Recruitment and Hiring. At that next level, there's a list of things that we would look for evidence of. When I looked at this and saw how this document had been pulled together, I said to some of my staff, "We are missing something here." We are missing actual staffing data and the results of the hiring. So that's an example of something we'll be adding. Then we'll identify sources of evidence. This is when the team sends out advanced notice to the district of certain things they'll submit, certain things they will just prepare and have available to the team when they come. But in other words, we are giving a lot of direction about where to look and what to look for and not just asking people to anecdotally and generally draw conclusions.

Another example is Professional Development, which is at that next level under the general Indicators. Over time, we'll be able to narrow the scope in these areas and say, "These are some of the most important Indicators, and these are some of the best sources of evidence, the places to look first, and then you only have to go to the secondary sources if you don't find the answer in those primary sources." Again, this is work that's being done and will happen over the next year. We'll be refining the process of how do you go out and look at all of this. Finally, I want to say a couple things about the piloting process.

We have a few people from the Department and a number of consultants working with us who are experienced school administrators, some are professional program evaluators by trade, some who are experienced public managers in other fields. And they are finding that this is a very useful process, and that we've had a very positive response from the districts where we've gone in to pilot this. The discussion back and forth is very useful. What was reported from the teams as a common response was how glad the districts are that somebody's coming out to find out what they are doing and not just making some judgment from afar. So I think that the process ultimately will be well received. I believe we can have a rigorous set of standards and pay attention to the work that districts are doing, commending those things that are positive while giving a push in areas where things are not going well. I welcome your discussion about the specifics and the substance of this.

**DR. THERNSTROM:** Juliane, I do appreciate very much the amount of work you've put in this and I'm delighted you have had a positive response from the districts, so far. But, alas, I do have some questions about this document. The main thing that's bothering me is that when I look at this list of Indicators under each category, I don't know what a lot of them mean. For instance, "Teachers in the district's schools use instructional methods and practices that meet the learning needs of their students." It's just too abstract, too vague. It's an invitation for very subjective, arbitrary judgments. "The district recruits, employs and retain qualified professional and support staff whose knowledge, skills, experience and personal attributes fulfill the district's personnel needs." Again, I don't know what that means.

"Teachers have high expectations for their own and their colleagues' skills and knowledge, and participate in professional activities in addition to the district's professional development offerings." Again, it's just too abstract, too vague. "Teachers and administrators in the district's schools have high expectations for student achievement and encourage student initiative and attention to quality." It's the same problem. I could keep going. If this is how we're going to evaluate districts, then I would say there is no argument against evaluating students with portfolios, because this is the equivalent. We have said we want objective measurements. That should apply throughout our system of accountability. These are not objective measurements. They are, as I said, an invitation to arbitrary, subjective judgments, and there's an awful lot of reliance from inputs rather than results. You know, my bottom line here is: simplify, simplify, and simplify. And, in addition, schools

educate kids, not districts. It seems to me, you want to know what the organization and effectiveness and efficiency of a district is, but it is at the school level that you want to look at what's going on in terms of curriculum, instruction, and so forth. And, again, you want to have measurements that are very objective, that are very measurable, and this just doesn't work for me. You know, again, I say this very reluctantly.

**MS. DOW:** I think that the specificity that you're feeling the need for, and I feel the need for it, too, is at that next level of evidence. What we have done is essentially taken what were standards and then indicators and then evidence, and the evidence is where you get at the more quantifiable data. Now that we have lifted it up to this next standard in order to group things and now that we have the focus on fewer things at the top level, I think that –

**DR. THERNSTROM:** But, Julianne, look. Let's take one of these. "Teachers in the district's schools use instructional methods and practices that meet the learning needs of their students." From effective school to effective school you can have very different instructional methods and practices. I don't want a checklist. I want to look at what the results are in terms of student performance, but I want to be able to say to schools, "Get there anyway you want, because there is no magic formula."

**MS. DOW:** But the specificity there would be what were the student results, what are the student results, and not just for students in general, but for all subgroups of students. So the first set of indicators that we would look at, or the first set of evidence that we would look, at are student results.

**DR. THERNSTROM:** Then school instruction comes down to one criterion: Are students learning by the single measure we have set up at the end of the day, which is the MCAS result.

**CHAIRMAN PEYSER:** Let's push that a little bit further, because, as you know, MCAS performance goes across a fairly large range, and we are looking not only at the range of performance, but also improvement over time.

**DR. THERNSTROM:** Right, but that should be school level data.

**CHAIRMAN PEYSER:** That is school level, but we also have an obligation statutorily and under regulation to make similar judgments about school districts, and we have the right to declare districts to be underperforming. We have two issues. One is we have districts at one end of the spectrum and at another end of the spectrum with, let's just use one measure, MCAS performance. You have most districts that are somewhere in between. I'm not sure you can develop any judgment about the quality of the district based on their performance, especially the vast majority in the middle. The other thing is that if we're going to make judgments, especially around underperformance, we have to do more than simply look at the MCAS data itself in the same way that the schools do. We look at things other than MCAS data. And we had long discussion about dropout rates and other things. But even beyond the data itself, schools and districts are complex organizations where, in fact, things may be changing and changing for the better, but it may not be reflected at the moment we are taking the snapshot in their MCAS performance.

**DR. THERNSTROM:** That's very true. On the other hand, what you're really saying to me is that, inevitably, there's a certain signal about district-wide performance here. You send a team in, and there are going to be a lot of judgment calls that are very subjective. Let's not pretend that we've got some kind of a list of objective criteria here that tells us anything. This is the kind of vague language that people in education retreat into because of the basic problem in pinning this stuff down. But my own preference is to not pretend we are pinning anything down.

**CHAIRMAN PEYSER:** I think the balance that we are trying to strike is narrowing the source, the points of the evaluation sufficiently. That way, people know, in general, what we're going to be looking for, while allowing for professional judgment on the part of the evaluators. Then, when they evaluate data, much of which may be conflicting, they'll have enough latitude to exercise judgment. I don't think you can make decisions, especially around underperformance, about the quality of these organizations, without having a team of trained, skilled, experienced evaluators exercising their professional judgment.

**DR. THERNSTROM:** I would be much happier, then, under Curriculum & Instruction to say that an evaluator should look at curriculum, instructional time, instructional practices, student assessment, and instructional materials and equipment. All the rest of this language I can't possibly figure out, including "the amount of time devoted to teaching and learning in the

core subjects is sufficient." I don't know what "sufficient" means. I have no idea what "sufficient" means, and it differs from school to school.

**CHAIRMAN PEYSER:** Under the Key Question around curriculum & instruction, let's say that the principle measure of the quality of the curriculum & instruction is their performance on the MCAS and other objective external measures. Let's say the school doesn't look good, the district doesn't look good, and the schools collectively don't look good. At that point, it's almost like a doctor trying to diagnose the disease. There are certain hypotheses or diagnoses that you go down as a potential cause for this problem: the curriculum; the amount of instructional time; the quality of instructional practices; the absence of other student assessments; or the absence of a system for using assessment data; inadequate materials. This is a useful checklist to go down and say: Is this the problem? Is this the problem? Is that the source of the problem? One of the reasons why the term Indicators is used as opposed to "standards" is to allow this kind of flexibility so that the evaluators can focus on those things that seem to be the most problematic.

**DR. THERNSTROM:** But at the end of the day if one district says, "These evaluators came to very subjective judgments, look at the district next to me which was also evaluated and things were weighed very differently." That district would have a very good case to make. These are too open-ended, too vague, and too abstract. They are not workable as stated.

**COMMISSIONER DRISCOLL:** Well, Abby, talk about coming in at the 11th hour.

**DR. THERNSTROM:** I have said this thing to you before.

**COMMISSIONER DRISCOLL:** We have talked about this for two years. We had representatives from Houston come to us in January of 1999. We have been talking about these kinds of standards around district evaluation, which is our responsibility under the law. I'm not going to disagree with that statement that it's really at the school level, but we are a system of 300 or so school districts. And it is the case that in other states, this does drive reform. So make suggestions on language or whatever, but we better be ready because the law requires us to be in the district evaluation business.

**DR. THERNSTROM:** I just made a suggestion that you just list these indicators without any further --

**COMMISSIONER DRISCOLL:** Let me try it my way and suggest it's not jargon. "Teachers in the district use instructional methods and practices that meet the learning needs of students." This says very real things to me.

**DR. THERNSTROM:** It will say different things to different evaluators.

**COMMISSIONER DRISCOLL:** But this whole system, which has been presented to you four or five or now six times, cut different ways depending on how the Board reacts, has a whole subsystem of evidence. We are not going to go in and say, "By the way, this just doesn't feel like you're meeting the students' needs." This is going to be based on data, on all kinds of and as much evidence as we can. You know, it's too early to say that District A is going to complain about District B. I don't believe that. I believe that we're going to go in and say, "Look, do you have a school improvement plan and do the goals of the school improvement plan meet the district's professional development plan? Are there instructional practices? Does it meet the needs of your students?"

**DR. THERNSTROM:** I don't know what that means.

**COMMISSIONER DRISCOLL:** Does it meet the needs of bilingual students by putting everybody in the same program? Do you have immersion programs? Do you have pilot programs? There are a whole series of ways to see whether or not a district has a plan for going about meeting the instructional needs of their students. This may not be perfect, and I will be glad to go through different terms, but every time we bring it, you want less. So we categorize it. I think we have learned a lot from going out to the seven districts, at least the people that we have had have learned a lot. I don't think there's the jargon there that you think is there. Plus, there is far more about outputs and results that get imbedded in all of this. I mean, this thing is this thick. That's what Julianne tried to say. That's only the front door. So do I think it's very steeped in evidence and will be very helpful to districts? I hope so.

**DR. GILL:** I'd like to say that I think that it's an excellent framework, and I do believe that the whole development of a performance measurement system, a performance standards system, is an ongoing process. You're not going to get it the

first time; nobody can, nobody ever has. And although I am coming in at the 11th hour and don't know all that comes behind this, I did hear your presentation in January, and I believe, as you said today, that there are measures behind it. And so what we have in front of us is an overview, it's a framework. I hope that we can get into discussion today in terms of what comes at the next level. But I would differ with some of the criticisms that were just made because I think it is a good framework. If you did not have your panel, I might have a few more concerns. But you do have a panel. The framework for your panel and the approach that the panel is going to take provides you with the kind of expertise and accountability that our schools need. If you have just objective data and you're looking at data points all the time, you're not going to have a system that gives you the kind of quality that I think we all want to see. You have to take too many things into account. Therefore, you need a process that will allow you to look at the whole. And I believe that this is what this does. I would like to hear more about what comes under this, and I specifically would like to learn a little bit more about when a school improvement plan is developed and the scheduling process, because I didn't have a clear understanding of how much time is allowed in order to really collect some of this information, to put the plans together, and how it fits, because there's a lot of work that has to be done. But I congratulate you for the work that you've done.

**MR. LaFLAMME:** Abby, there's some fuzziness here, you might say. I think you could make a decent argument that there's some fuzziness implicit to this whole thing. But is the more desirable route to put together an equation into which we check things off based on quantifiable indicators?

**DR. THERNSTROM:** No. I would just like to get away from language I don't understand.

**MR. LaFLAMME:** Right. But it seems that your objection stems from the fact that it's not quantifiable.

**DR. THERNSTROM:** No, I literally have no idea what most of this means. I just gave a few examples. I can give a lot more. I don't know what it means.

**DR. GILL:** Won't we learn some of what it means when we talk about the next level, Juliane?

**DR. THERNSTROM:** Just take the first one. This is just arbitrary. "The amount of time devoted to teaching and learning in the core subjects is sufficient." What does that mean exactly? From school to school? I mean, there is no amount of time sufficient that anybody can come up with as one person's amount of time "that meet the learning needs". I don't know what that means.

**DR. SCHAEFER:** To try to be more concrete about what Abby is saying, take that first Key Question, "Of sufficient quality to meet the educational needs of all the students served." If there are students who are failing MCAS, does that mean that the district has not provided sufficient quality in terms of its teaching for all the students, to meet all the students' needs? Obviously, the goal is for everyone to pass. So are we condemning the district because X percent are failing? It's not clear from the very beginning, from that question.

**COMMISSIONER DRISCOLL:** The Key Question is meant to be the overriding question. I'm not going to worry about the word "sufficient," to get back to Marcel's point, or that we have to have 847 minutes in mathematics. But I know a high school that recently adopted a block schedule as their way to meet Time in Learning. In their iteration of establishing block scheduling, there is a certain period when sophomores don't get mathematics at all for about a six-month period. It's a problem. I would want to know what the district is doing about that. Today's kids are bombarded all the time by the media. There's enough regression that takes place over the summer and the April vacation. We either have to talk about the specifics, which I'm glad to do, or we talk about the totality, but we've done both. In other words, the whole issue is a system whereby this is supposed to hang together. The question is backed up by these headings which are also backed up by a whole series of evidence. I know what an Instructional Practice is; you know what an Instructional Practice is, so the question is are they doing it just because they've done it since 1968? Or how are they organized? Are they heterogeneously organized or homogeneously organized, and why? What kind of series do they have? You simply want to know what their Instructional Practices are and if it ties together. I don't think it's that fuzzy, although it's not as specific as to say it is demonstrable in numbers.

**DR. THERNSTROM:** You tell me what "high expectations" mean. One person's high expectations are another person's too low expectations.

**COMMISSIONER DRISCOLL:** When I walk into a district I know whether their practices are built around the fact that kids can learn as opposed to house them, feed them, and get them out of here. There's a huge difference. As for the words "high expectations," what are you going to use? "Moderate expectations"? I'm not going to defend every word of the language. It could be crisper, I suppose. But the issue is to have a set of standards by which we could look at a district and see what they are doing.

**MS. CRUTCHFIELD:** David, can I jump in here and ask what the definition, specifically, of "high expectations" would be as you walk in the door?

**COMMISSIONER DRISCOLL:** Well, again, we're talking about two words out of one standard under a question.

**DR. THERNSTROM:** No, it comes up again and again.

**MS. CRUTCHFIELD:** I just wonder about the general language and I wonder if there are concrete and specific guidelines...

**COMMISSIONER DRISCOLL:** Let me ask it this way.

**MS. CRUTCHFIELD:** When I'm finished-- measurements by which we make those general words concrete.

**COMMISSIONER DRISCOLL:** I had the pleasure of visiting five schools in Massachusetts with Bill Edgerly, who donated \$50,000, \$10,000 each for the five principal schools whose MCAS scores were the highest in the state. One of those schools was in Springfield, one of those schools was in Boston, one of those schools was in Revere, and we took a video camera, we met not only with the students, but we met with the teachers and, naturally, the principal. We asked, "Why do you think your school did so well?" Every single one of those five schools talked about high expectations.

**MS. CRUTCHFIELD:** How did they define the high expectations?

**COMMISSIONER DRISCOLL:** The fact that they don't settle. That's the best way I can say it. We can nit-pick all day about defining "high expectations". I know when I walk in. You know it when you walk into a gym whether the coach has high expectations. You know whether the school has high expectations for its students. Now, there's certain evidence that will come within it. So we are either going to argue over the words that we use as part of the standards, like "high expectations" or we're going to not appreciate the evidence that it will lead to.

**MS. CRUTCHFIELD:** What does the evidence look like?

**MS. DOW:** Let me talk a little bit about this, because this is one of the interesting ones. When you're talking about expectations, you're asking how we get at people's perception of what's expected of them? In the state of Rhode Island and the Carnegie Middle Schools Project they are using a surveying approach where they ask a set of questions of students, parents and staff about their expectations given the amount and quality of work that's performed, and given plans after graduation. There's a set of questions that are posed and data that's collected around what those expectations are. One part of it is the observation of rigor in the classroom, and looking at student work and evaluating the quantity and quality of work that's expected of students. Part of it also is general perceptions about what the students are capable of. We have done some surveying in preparation for the school panel review visits which are coming up, and some of the questions we've asked have to do with teachers' perceptions of the barriers for students learning and do they expect them to do better on the next year's MCAS tests. That's telling. The responses in some places are that they don't have high expectation that the outcomes for students can or will improve.

**CHAIRMAN PEYSER:** There are some other things that would emerge, especially over time, as more of these are done. For example, going into an Algebra I course in one district and an Algebra I course in a different district, to see what kind of material is being covered at what level, what kind of work is expected of students, what kind of grades students get for the same quality work, and looking at MCAS performance or other standardized measures in comparison to grades provided in the classroom. I mean, those also tell you about level of expectations.

**COMMISSIONER DRISCOLL:** Seeing what kind of practices they are using and seeing whether it's fitting the needs of the individual students.

**DR. SCHAEFER:** I know I asked this question before, but I want to ask it again to make sure that the answer hasn't changed. Namely, are there any other states that have already instituted an accountability system?

**CHAIRMAN PEYSER:** District accountability.

**MS. DOW:** Nothing that looks quite like this. In the places we have seen that are doing a district system, some of them have a district accreditation process that is more based on inputs, less of a focus on looking at outcomes.

**DR. SCHAEFER:** So there's really no state that has a system that is trying to base it on outcomes at this particular point in time.

**MS. DOW:** Not that I've seen, not with a district performance process. Texas aggregates data up to look at the district level and they use data in addition to the testing system they use, but they don't have an on-site process. They haven't combined the two. So what we tend to have are states that have on-site processes used for accreditation, and we have other places that have used an outcomes measure, at the district level, that's based just on data.

**CHAIRMAN PEYSER:** Another thing to add is that there are evaluations of school districts on the basis of their compliance with federal and state mandates. One of the things that this attempts to do, subtly, is reorient the nature of accountability to one that is less focused on simply compliance and more on outcomes.

**DR. SCHAEFER:** So we need to be clear that we're talking about new territory here and this is an experiment. I share some of Abby's concerns. I'm also concerned about the extent of the evaluation. I understand that there's going to be data collection before the actual visit, but I am concerned about how long each one will take and what is the expectation. I know what we've said about the amount of time that will be spent at the district itself being three, four days. But what about the time that will be needed to absorb all the data that you get before the actual on-site visit? What will that mean in terms of the cycle that we have talked about, that is, how many districts we think we can do each year? When I look at this, frankly, I am overwhelmed. I can't imagine the staff you will need to get through even 10 or 20 in a year.

**COMMISSIONER DRISCOLL:** We are doing 60.

**DR. SCHAEFER:** That's what you're saying, but you haven't done that yet. That's what your expectation is.

**COMMISSIONER DRISCOLL:** You're absolutely right.

**DR. SCHAEFER:** I'm asking whether that's realistic. Have you thought through this data collection phase? Have you thought about how much time it will take to actually compile it and absorb it? Do you have a good sense that all the people who are going into this district have a sense of what they are looking at before they go in? Do they have this data either on paper or in their heads or something? I'm really concerned about that.

**COMMISSIONER DRISCOLL:** I share that. I think it's a colossal undertaking, no question about it.

**DR. SCHAEFER:** Maybe that addresses what Abby is saying about the extent of the questions and about the need for simplification in a certain respect in order to do it in a reasonable period of time. At the end of the year you don't want to have done 10 districts when you intended to do 60.

**CHAIRMAN PEYSER:** I think our issue at the outset needs to be honing the process and improving the quality. If that means we have to sacrifice on the quality, I think that's the tradeoff we have to make.

**DR. SCHAEFER:** I think we need to be clear.



**COMMISSIONER DRISCOLL:** You're absolutely right. A lot of this is going to be learned as we go. I believe that dynamic is going to cause a lot of positive results in and of itself. We are not going in saying that we have the answers; we're going in saying that we have a system, we think it works, we're going to do it, and learn from you. That's why we tried the seven districts. This first year is going to be a tremendous learning process.

**DR. SCHAEFER:** Let me ask one other question. Is this system that we are establishing for districts more similar or less similar to the evaluation that's being used for the charter schools?

**CHAIRMAN PEYSER:** I think, and Juliane can talk more about it, it's different. I think there are some things that may be similar in terms of process and in terms of the overall approach but in my view, an evaluation of a district has a lot more emphasis on systems and organizational structure and leadership than you will find at a school level. Those things are important but you're probably going to be focusing more time on classroom practices, teaching and learning.

**DR. SCHAEFER:** Let me ask this: For the district, then, when you go in, you're obviously going to have to select a couple of schools. You can't possibly cover -- if you're talking about Boston or Lowell or Worcester or something, you can't possibly be looking at all the schools. So you're going to have to select from among those. I assume you're going to select a cross-section of them so that you're running the gamut of schools that are doing well on MCAS, schools that are doing poorly, and in the middle, and that in itself is going to take an incredible amount of time if you're going to be fair to those schools and that district.

**MS. DOW:** Again, we talked about that. In our large city districts we would definitely be spending more time. It's not something we would do in five days. And we'd also have a much larger team. But we only have a few of those large districts. Most of our districts are small or of a manageable size for us to do in this other model.

**DR. SCHAEFER:** I think we have enough districts that are certainly multi-school. Most of our districts are not just one school.

**MS. DOW:** It's not that you would visit every classroom. There is something different about doing a school-level inspection. I'm not sure if they've done this in the charter school, but they have done this in the British model. They have gone to every classroom and done an observation more than once in every classroom, and they are able to talk program by program and give individual feedback to each teacher. We would not do that at the level of the district review, although good practice might suggest that the district would do that periodically in each of their schools. In fact, Boston, Springfield, and, I think, Worcester has begun a process to systematically evaluate their schools. I think that's the appropriate relationship for us to have-- encouraging a process. Now, the larger the district, the more systems they'll need in place to do that. In smaller districts they may do it in a much less formal way, because the superintendent may visit every one of the schools every week and spend a day there. In a small district, they may have much more intimate knowledge of what is going on at the teaching and learning level.

**DR. GILL:** As I listen to you talk, you keep referring to the fact that we're going to learn something and it is iterative. It's also very much interdependent; there's no one indicator that's going to be used to identify a district as underperforming. I do understand, notwithstanding my initial comments, the concern that Abby has. I think in part that comes from using the word "standard". When we see the word "standard", we think of something as very specific. And I'm not familiar, David, with the legislation and any requirements that you have for using the word "standard". Is it possible to use "district performance guidelines", "measures", to use something other than "standard"?

**COMMISSIONER DRISCOLL:** Obviously, you have to have standards. But we can call the those indicators or --

**DR. GILL:** Abby, I don't know whether that might help you, because when I see the word "standards", I think of something as very specific.

**DR. TERNSTROM:** Look, what's bothering me, Dave, is the answer, "You know good education when you see it." It's equivalent to the Supreme Court's test for pornography-- "You know it when you see it." And when Jim says, "Well, when we are looking at, for instance --

**CHAIRMAN PEYSER:** I don't think I used that tone of voice, Abby.

**DR. THERNSTROM:** I'll change my tone of voice. Jim, when you were talking about learning environment and your translation of high expectations, you said, "You can look at whether the courses, like a math course, are actually rigorous." Well, okay. We can look at an algebra course and see whether it's back-of-the-bus algebra or a truly good algebra course. But that is a school-by-school curriculum question. And Dave, you've said, regarding instructional practices, that you can see whether they are heterogeneous or homogeneous groups but good schools have a variety of ways of organizing kids. You gave the example of bilingual ed. Well, there's an ongoing debate over how best to deal with limited English-speaking students. My point here is, since the basic test is really, as you describe it, knowing good education when you see it, why not admit it and simplify this? I don't mind having something that says Curriculum & Instruction and just list curriculum, instructional time, instructional practice, and student assessment. Don't try to define any of those.

**COMMISSIONER DRISCOLL:** I will try this again, Abby. It is a district and there are various pieces of evidence and there also are observations. I was giving examples, not trying to give the only example. Does the district organize at the elementary level heterogeneously because it has a reason for doing so? That's the whole point. It isn't a question of taking the one indicator.

I didn't say high expectations are about whether you see it. I know a lot of schools that think they are wonderful. I mean, that's the great story of Americans. American kids take the test and they think they did well. They didn't, but they think they did well. So the Japanese study us because they want to know how we get this great self-perception that our students seem to have.

Why do you use the term high expectations, what does it mean? I mentioned five schools that used that term. And it's not just about looking at the algebra course. Just because the algebra course has high expectations based on high level of content, it doesn't mean it's being taught well. So everything is interlocking. I don't want to simplify it; if I did, it was because I was answering your specific issues. I think it's wrong to limit things to one statement because that just doesn't do justice. And I understand the problems of having 76 standards, but that's the balance we have to come up with. Some of them are indicators and you go look for other things based on the evidence.

**MR. LaFLAMME:** Can I ask a quick question? When we say, under Instructional Time, "the amount of time is sufficient for teachers to provide direct instruction," if we list that and hand it to our district teams and suggests they look at that, aren't they going to be asking themselves, "Is this sufficient?"

**DR. THERNSTROM:** Marcel, you tell me what, "Teachers in the district's schools use instructional methods and practices that meet the learning needs of their students" means. Translate that for me. I don't know what it means.

**MS. CRUTCHFIELD:** I think it's translated into being there.

**DR. THERNSTROM:** That's my whole point, that this is a "you know when you see it".

**CHAIRMAN PEYSER:** Let's say you're talking about limited English-speaking students as being in one category of need, kind of an identifiable category of learning needs. An overall view of the data of the district may show that students are performing well in English, but in disaggregating the data you may find that English limited students are performing poorly even relative to other similarly situated districts. That would then lead you to investigate further asking, "What are they doing? What is their approach to teaching limited English-speaking students to speak and read English?" Then, you make some evaluation, again on the assumption you have informed people who know and understand what they are seeing, and know how to evaluate the information that's available, who will make some judgment about whether they are really not meeting the needs of these particular students.

**MS. CRUTCHFIELD:** I will tell you what my response to that is, Jim. We would be much better off, under Curriculum & Instruction, doing a listing. Evaluators are going to look at Curriculum, Instructional Time, Instructional Practices, and then giving very concrete examples, like the one you just gave. For instance, they may see in a district that the limited-English students are underperforming relative to what any kind of reasonable expectations are, and therefore the evaluators should be zeroing in on that. Or your previous example on an algebra course, which is very concrete. We could give a series of very concrete examples for evaluators, and then you would have some real guidelines. This kind of language doesn't give any guidance as far as I'm concerned.

**CHAIRMAN PEYSER:** We definitely should do that. Some of that might already exist since you're not seeing the whole document. But I think the point is well taken. If you're going to train people to be evaluators, you have to do more than just show them words and hope that they interpret it in a consistent and effective way. You have to have people with experience. You have to go through the process of explaining what is meant by this and, in particular, how they are expected to exercise informed judgment using facts and information that may or may not be readily available.

**MS. CRUTCHFIELD:** I would hope that that's the case, and I'd also hope that it's the case that we provide the same kind of material to the schools or the district that we're going to use so we have a common denominator. I think that while this is a work in progress, my assumption is that there's a structure, a system in place to make for common language and understanding: collect data, review, plan, do, check, act, do a quality search, et cetera.

**MS. DOW:** The pilots are serving to do those. What they are doing in the pilots is coming out with those questions. They have a list of evidence underneath that. It's really a prompt for the evaluators to think about this and that, and be sure to look at the different categories of students within the system. And so they actually have something that becomes their checklist. That's what that next layer becomes, evidence to look at. But in the pilots, they have said, to focus groups of teachers or administrators, "Looking at this standard, what evidence would you offer us? How would you answer this question for your district? Are you able to answer it or does it vary school to school?" Some of them have said, particularly in the areas that relate to Curriculum & Instruction, that they really can't give us a district-wide answer to this because they don't know, at the district level, what's going on school by school. They don't have any systematic way of gathering that information right now, but maybe they should.

In other areas they are able to show us things at the district level. They begin to talk about and pull out from the documents that they've assembled some things that show the plans, the results, their staffing, their processes, their communications with parents, whatever is relevant to the topic. Then the evaluators take their list and ask, "What about this? Do you think that this is also significant or do you have any evidence that relates to this?" They are using their list and they have an exchange of that sort. From that we think we'll be able to hone some common agreement about things you should always look for and ask about. For other things, we want to leave the option that we would ask more open-ended questions such as -- Are there other things that you would show us? It doesn't become just a checklist and something that is closed ended.

**DR. SCHAEFER:** This is a different piece of it, but under the people who are going to be parts of the panel, I'd like to see just that each panel include a non-educator.

**CHAIRMAN PEYSER:** Are you now talking about the school review panels or the district evaluation teams?

**MS. DOW:** The district has evaluation teams and, yes, it is contemplated that at least one member of the team would be somebody coming out of business or something other than the public education sector. Even on the pilot teams that we have out, we have people who are --

**CHAIRMAN PEYSER:** Let me ask a question which relates to the issue of why the word "standard" is there, because it is quite possible that a district goes through this process and comes out as being underperforming which has certain consequences attached to it. How would this process, or how would the groups conducting the process reach a conclusion of underperformance?

**MS. DOW:** Ultimately we anticipate that at the end of each of these indicators there would be a rubric and there would be levels of performance, one of which would be unacceptable practice. These all represent things that could be, of which there's a range of performance possible from purely unacceptable to excellent performance. By going out and piloting it we are developing just that. They'll be continuing to do it next fall as we roll this out. It's further assembling the samples of work, if you will, that's parallel to what we have done on the student side. It involves looking at the range of practice and what we see when we investigate an area. What does excellent practice look like and can we have some common agreement of it? We did some early work on this saying what those rubrics look like, and they were quite mechanical in their construct and didn't really describe anybody's real practice anywhere. The real practice is going to be complex. One of the things that we'll see as people are doing these evaluations is what comes across as being really valued and valuable and therefore might push something up toward the higher standard.

**CHAIRMAN PEYSER:** I understand the need for having rubrics in order to allow for consistent evaluation, but I'm concerned about that becoming a straight jacket whereby the sum of the parts doesn't actually add up to what is perceived to be the whole. What I mean by that is the answers to the six Key Questions. It might be that in practice, because of the mechanics of the process, you go through the individual pieces, apply some rubric and, gee, it doesn't look like they are doing so well. But when you roll it up, they are a high-performing district. When you look at the output measures, they are a high-performing district. This relates to the issue we were talking about earlier: How do you know what you're seeing? How do you evaluate it? You need some more texture and some more detail around exactly how to guide evaluators. You don't want to make the individual judgments on these indicators a straight jacket that prevents them from making what to them may be the obvious conclusion that this is a district that's in trouble or a district that's excelling. Though that's the first thing.

The second thing is the programs that are related to that. I'm picturing a district which, for some reason, seems to be performing well on external measures but for another reason appears to be disorganized and sloppy. And it may be one of these things where having a disorganized district has turned out to be an asset for the schools. I don't know. When you run the math they are performing well, their superintendents are learning, but when you add up all the numbers, they look like they are an underperforming district.

**MS. DOW:** I think this comes down to why we declare districts underperforming. If we have determined they are underperforming, we have determined that they are not doing the job of adequately educating children in the district; therefore, the state must do that job where we have a constitutional obligation. In the case you just described, where student results are acceptable to everyone, but we find the management is poor, should the state put that district in receivership? I don't think so. I think there are going to be other places where we need to go first. And that's exactly why we're giving that the weight.

If we find, on the other hand, that the results for students are poor and that they are not improving, we go in and look at the quality of the curriculum and instruction available to students. Then, if we look at the organizational processes that should be supporting a good quality education, and it is not there, that's when we'd be finding unacceptable organizational structures and a consequence.

**CHAIRMAN PEYSER:** It's a two-step process in terms of making a recommendation of underperformance to the Commission or to the Board. The first step is to look at the MCAS data of the schools in the district and to make some judgment about whether the district is meeting expectations and whether it is at an adequate level. If the answer is no, then we are going to do the evaluation. In terms of making the decision, we will be monitoring the evaluation process to determine whether or not the data is capturing the true picture or is a snapshot in time, which has changed or is changing.

We need to think a little bit more concretely about the decision-making process around underperformance. It's probably going to be something like a two-stage process similar to what we do with the schools. First, we'll look at the output data of student performance. If that's inadequate, then they are eligible for underperformance and we will do the evaluation. However, we are not concluding that they are an underperforming district.

**DR. THERNSTROM:** After you look at the output data, Jim, you still have the central question that you posed, which is: How do you know what you're looking at? And the vagueness of a lot of this language, it seems to me, invites districts simply to get onto the right rhetoric. Oh, yes, we have high expectations for our students. Yes, we have a good reason for heterogeneous groupings, or whatever. That's quickly learned.

**CHAIRMAN PEYSER:** Actually, it's a timely comment because Dave and I and Julianne were actually at an EMAB meeting with Mark Roosevelt where he commented that it takes about four questions to cut through the prepared answers to see if there's anything there. This is why you have to send people out there. You can't rely on written reports or submissions or questionnaires. You have to go out and look someone in the face and ask the questions to see if there's anything behind the first level answer. If you've got the right people, and that is all dependent -- this whole thing is dependent on having well trained, well-qualified people. If you've got those people applying something like this, I think you can cut through that.

**COMMISSIONER DRISCOLL:** I think the examples that Julianne gave on the issue of high expectations are examples of where this system is not just coming in the front door. This has deep roots in terms of the kind of evidence and rubrics that follow underneath it. We are not going out and saying, "Hey, by the way, do you have high expectations?" And them

saying, "Yes, we have high expectations," and then we move on. I don't know what to say, Mr. Chairman. I think we may need to take some time and again go through the entire system.

**CHAIRMAN PEYSER:** It's 5:30 and we were supposed to spend the minority of our time on districts and the majority on schools. As you know, there are eight middle schools in less than two weeks about to be visited by state inspectors. It's important for us to understand, before that happens, what the basic design and process is and what feedback that might deliver to us as a Board in terms of the declaration of schools as being underperforming. So what I'd like to do is shift gears to that. We do have this on the agenda tomorrow for further discussion, and maybe between now and then we can talk about any specific recommendations we might make to the Board for next steps on this. But I think if we can shift gears and go to the schools that would be helpful.

**MS. DOW:** We're going to eight schools. We have panels. I wanted to just give you a sense of the panel composition, among other things, because that's one thing I didn't put in the packet. I don't know if you had a chance to read what I sent out that describes the whole process that we'll be going through.

We have four panels; each panel will be reviewing two schools. The panels are made up in large part of practitioners: school principals from urban districts, curriculum coordinators from both urban and nonurban districts, and then each panel has at least one person who's been involved in doing charter school reviews or other kind of school evaluations. There's a facilitator who is either a member of the Department staff or a consultant. In one case, there is a person from a district whom we've asked to play that role. In addition to having these people, we have two other people for each panel. One is a panel monitor whose job is to watch how this whole process plays out. Because we are trying out a process which undoubtedly will be imperfect, we're going to be studying what we are doing while having an evaluation process. And we've contracted with School Works, who is the contractor doing our charter school reviews, to develop an evaluation protocol for us and to coordinate that evaluation piece. We also have four individuals who were from the Department who are serving as the logistical coordinators between the Department and the schools.

They'll carry over and be the link into the targeted assistance that will be provided at the other end of this process. So we are using this as a way for Department staff to gather all this information and to provide a helping role, but those individuals won't be part of the discussion or decision-making process that goes on during the panel review. They will be observers and logistical support during that process.

We have people involved from Lowell, Everett, Framingham, Boston, Brockton, Lynnfield, Marlboro, Springfield and Worcester. We have a nice array of communities represented. We made an effort to reach out to practitioners in those communities that are likely to be affected by this system in future years, if they are not involved already. We invited practitioners who were in schools where there has been improvement on MCAS. We already had a couple sabbatical teachers, curriculum coordinators, and school administrators.

The process involves the team members first getting together for a four-hour session where they have the data packets presented to them. Then, they take the data and begin to work with it as a group. That happens this Thursday. Next week, the first two panels will be in their districts, with two more panels the following week. Contrary to what was in the paper today about there being a half an hour of consideration, these panel members will spend four hours in this orientation and then they will take this book full of information, from the Department and the schools, and study that. This book includes answers to preliminary questions. They will return to spend four more hours, the day before they go out for their review, and then spend ten hours on the actual review day focused on that school.

The two questions that the panel is being asked to advise the Commissioner on are: (1) Is there a sound plan in place for the school's improvement? and (2) Do conditions exist for its successful implementation? Those questions have been posed to them. There are also some subsidiary questions that have been provided to them to guide them in looking through that. So they'll do their own analysis of the data so they can determine whether or not they think the plans are sound from the perspective of what they understand about the school and its needs from that. They'll then speak with the leadership at the school and see what their analysis is of the need and how it's the same or different from what the panel had come up with. They'll look at the plans in light of that. They'll interview staff as well as meet with the principal and the leadership team at the school. During the morning and afternoon session, they'll have an opportunity to meet with representative of the school committee, the superintendent, and the union. The responsible parties will be at the table so there can be a frank discussion about what the school's performance is and what the prospects are for improvement. The panels, in reporting back, don't

have to come to a unanimous view on the questions. There can be multiple perspectives reporting back to the Commissioner. Finally, the Commissioner will take all of the information, review it, and make a determination about whether any of these schools should be declared underperforming.

**CHAIRMAN PEYSER:** There's one piece of this that needs to be explicitly included, presumably under Guiding Questions for Review Panelists. It is evaluation based on other data whether the MCAS data is representative of the actual student performance in the school. I think there are two issues here. One is, explicitly looking at dropout and attendance data. Granted, these are middle schools, so the dropout rate may be less of an issue, but we need to take that into account. And the second is, especially on the standardized test data that's available, if it shows that last year's fourth grade class was just an anomaly, that's important to evaluate as well. As long as that question doesn't preclude going forward with the rest of the questions. It's important to verify the MCAS data up front as representative of the actual student performance of the school.

**MS. DOW:** Before I came out today I was talking with Paula Girard who is working at putting together the data packets for her team and working on the set of questions that relate specifically to the data. That was one: Are there any indications of improvement when you look at a more detailed assessment of the data?

**DR. THERNSTROM:** Juliane, I wish the district accountability system looked like this because I think this is great. School performance is the central question and I'm very, very struck by the simplicity and soundness of the description here of both the criteria and the process. And it does seem to me that if this is what the district accountability system looked like we'd be in great shape.

**COMMISSIONER DRISCOLL:** I'll tell you what. I will let you put a panel together and you go out with your standards, we'll pay for it, and we'll go with ours, and then we'll match it. We'll meet at noon.

**CHAIRMAN PEYSER:** I recently spoke with someone who had a similar critique of the district protocol versus the school protocol. If you look at the description of the school and district accountability system framework, there's a section with respect to evaluating district performance, which is not identical to the Key Questions here. You might want to look at that because there are actually 12 Key Questions here with respect to the districts. What is here is intended to inform and provide some additional texture to the six Key Questions here. I think there may be work we can do to try to bring these two things together. If this has greater clarity, if they are posing questions in this document that are essentially trying to get at the same point but do it in a more clear, concise way, then maybe we should be borrowing more from this.

**DR. SCHAEFER:** Can you make copies of this for us?

**MR. LaFLAMME:** It's a different scope of inquiry, as well.

**DR. THERNSTROM:** It is a different scope, but look, school performance is the bottom line. In evaluating schools, we have not gone with these pages and pages of vague criteria.

**MS. DOW:** Well, Abby, I just want to warn you that what comes after this is the set of criteria for actually going and doing the in-depth evaluation of the school that's declared to be underperforming. It's at that point where we will get into a protocol where we are doing a diagnostic evaluation. At the district level, we are not trying to diagnose and prescribe treatment where we know that there is a serious problem. This is really about whether or not we go to that next stage. I appreciate that it looks simple, but it's not. I assure you that when we get to the point where we are ready to go out -- as we are with these panels next week -- it will have that same clarity and it will satisfy you once we are able to articulate it. We'll keep working at it until it does, really. Not just you, because if it doesn't, it's not meaningful, as it's phrased. It can be rephrased because there is meaning that we have in mind behind each of these things that we need to go and look at. This has been hard, particularly that portion on Curriculum & Instruction and looking at what's going on at the school level. It's very complex.

**MS. CRUTCHFIELD:** I have a question about diagnosing schools that are doing things well. Where is the appreciative inquiry here? In your conversations with schools that are not doing things well, what are they doing well? What are they getting right and how can they build on that? For me the way to make things concrete is to help people become able to see

what a best practice is, what a benchmark is. So I think we are missing a big opportunity here if we don't have something in place that helps us diagnose and make meaning out of what is going well.

**CHAIRMAN PEYSER:** We are in a unique time frame here. Under the regs that we adopted, we authorized the Department to go ahead and do evaluations of the lowest-performing schools that appear to be showing no improvement. We're doing this before we have the two years of data that is supposed to be the standard process for judging the performance of districts. This fall, we will have two years' worth of data, past the baseline, that will allow us to run the system as intended. I think you're right. We need to look at both ends of the spectrum including what's in the middle, if for no other reason than to provide context for people looking to understand where this fits in the range of actual performance. That's a good point. Do you have anything to add, Julianne?

**MS. DOW:** Only that we're going to look at, after we complete this panel review process, those schools that show improvement. We're going to pick a sample and ask them to do the surveys that we used for the eight schools under review. We're going to compile the same data packet and begin to study what seems to be similar and different. Then we will be headed toward a process. Either we'll be able to use this same panel review process to look for those exemplars or we'll have a modified process that asks somewhat different questions for schools that seem to be improving.

**MR. LaFLAMME:** I love it. It's good. Just one thing, and it's a small thing in terms of the format of the document, I suppose, but I feel like it's a big thing in terms of what it implies. At the back of the packet, the Guiding Questions for Review Panelists, I think that's a good breakdown of the two Key Questions. Buy-in is an important category. "Are faculty and staff optimistic that their students' performance on this year's MCAS tests will be better than the results in 1998 and 1999? Do faculty have confidence in the school's leadership?" Given the idea that, referring back to the Review Panel Code of Conduct, students come first and are the center of review, I would love to see added to that category: "Do students expect that their own performance will be better? Do they have confidence in their school leadership?" That's key. I think there's a lot to be said for asking kids if they feel ready for these tests? And, if not, tell us why.

**MS. DOW:** We had some discussion about whether we could try to do data collection about student opinions and views at this stage and decided, at least for this year, that it becomes somewhat more complicated. We talked about both parents and students, but it becomes complicated because you need to be prepared to translate all of your materials into multiple languages to make it available to all students and families. There are some issues like that that were just logistical complications. But as we discussed it, we concluded that we should focus that inquiry at stage 3 when we do the fact-finding process. Then, there would be time to focus on interviewing students and observing with an eye toward trying to assess what's going on in the classroom. We also thought of doing some focus groups if not surveys at that stage. But we are trying to do a lot at stage 2.

**CHAIRMAN PEYSER:** Just to underline the point, once schools are referred for review to the panel, it's sort of like being in limbo. There's a certain time sensitivity in fairness to all involved to try to expedite the process, not to rush to judgment, but to make sure that there's some initial informed judgment made before the more in-depth review happens. It is worth pursuing student and parent questionnaires or surveys, in part because in some cases there may be some huge problem, sort of the elephant in the living room that everyone knows about but no one talks about, at least, unless you ask them. And sometimes it may be the only people you can ask that will say it are the students and the parents.

**DR. SCHAEFER:** You mentioned that each panel would include somebody who had been reviewing a charter school. And I think this would also be a good opportunity to try to get some data comparing the charter schools with the regular public schools to see in their observation what's different.

**MS. DOW:** Again, the scope of this review is substantially different than what they are doing for the charter schools. Certainly we'll be having some discussions. They may or may not feel that they are able to compare.

**CHAIRMAN PEYSER:** In the ideal world these two things start to converge. There are certain things about charter renewal and evaluation that are somewhat unique specifically when the charter document within the school is compared, but in a lot of districts there's similarity. So I think that it's important that we try to integrate the processes as much as possible so we can make some observations and draw some tentative comparisons and conclusions.

**MS. CRUTCHFIELD:** I worry about not including students in this round. I understand why not, but I worry about it because they are the bottom line. They are important stakeholders as such. I wish there would be a way that it could happen. What if you collect data from the adults and then add the students later? It is true that any time you leave out a major population in a system, and it might feel like a one-up/one-down, the feedback that you get may be different because they weren't included. Is there some way to reconfigure this so we don't leave students out and we don't make them wait until another level?

**CHAIRMAN PEYSER:** My only suggestion, the schedule is tight.

**MS. CRUTCHFIELD:** I understand.

**CHAIRMAN PEYSER:** But it might be organizing a focus group of students and focus group of parents or something like that.

**MS. CRUTCHFIELD:** That would be great. Something so that they are in this circle from the beginning and perhaps their involvement simply increases but doesn't go from zero to something later.

**MS. DOW:** Again, we talked about the idea of getting students together. We talked about the possibility of having a structured engagement with the teacher and students rather than doing classroom observation.

**COMMISSIONER DRISCOLL:** We have to understand that these eight schools are being referred for review. This is a big deal; they are under pressure-- which to some extent we want, to some extent we don't want. Just because a school has low scores does not make it an underperforming school. These are all urban middle schools and we are trying to process them for the first time. As Julianne said, if we get to a fact-finding stage, obviously, that is when we will interview parents. We have to learn from these eight. It's a big process.

**CHAIRMAN PEYSER:** But maybe one of the post-review questions is: How could we engage parents and students.

**COMMISSIONER DRISCOLL:** That very well could work.

**MS. CRUTCHFIELD:** One of the things we are perhaps learning already is that we have to find a way to invite parents and students to the table. I understand exactly what you're saying, but I'm clear about this.

**MR. LaFLAMME:** The logistics of a one-hour focus group are significantly less complicated than that of this widespread survey, et cetera. The idea of getting 15 students in a room with two of the panelists for an hour is tenable.

**MS. CRUTCHFIELD:** And rich with dignity.

**COMMISSIONER DRISCOLL:** Rich with a lot of things, data being one of them.

**MS. CRUTCHFIELD:** That's right.

**CHAIRMAN PEYSER:** Any other questions or comments? Somehow we managed to end early, which seems impossible given how we started.

**COMMISSIONER DRISCOLL:** In the two minutes we have remaining, I just want to remind the Board that this is a process that began almost two years ago. I'm expecting to get a vote tomorrow, although we did build in a period of time for Board members to provide additional information. At least that's my recommendation. But we do want to send this out for public comment and I hope the Board will agree to do that. I'd be more than happy to arrange the time to have Board members come in, but we want to get this out for public comment which, by the way, gives us all time to continue to work on it over the next six or eight weeks. In fairness to the districts, we want to have this back for a vote in June so districts know exactly what it is that we're going to be using as an instrument for the next year.

**MS. DOW:** That's a perfect segue, if I may, to just two endpoints of information. One, we already have some superintendents who have agreed to host some roundtable discussions about this at the end of May. There are districts that



have participated in the pilot, either by being a pilot site or by being on a piloting team. So, again, that's a way of getting the field engaged in the discussion. And also, the reason why it's important for us to have a set of standards and indicators adopted, at least for what we're going to use for next year, is so that we can prepare the materials to send out to the districts that will be reviewed in the fall and to prepare for training of the evaluators that we'll be bringing on. And we have entered into a contract with a British consulting firm to work with us on the development of the training. So we have to know what we're going to be training on, too.

**DR. SCHAEFER:** Jim, you mentioned being at a meeting with the EMAB Board. What is the relationship between what we are doing here and that group?

**CHAIRMAN PEYSER:** This is an open question at the moment. Let me preface it by making another comment. Part of the structure of our accountability operation is the creation of an advisory panel that would be staffed by the Deputy Commissioner for Academic Affairs. That has not been done yet, in part because of the second point, which is that we are in limbo in terms of whether the legislature will enact something that causes us to have to reorganize the office. Specifically, there are proposals to either require collaboration with the Department of Revenue, which is going to happen anyway whether it's statute or not, and in a different proposal, to create an independent office nominally within the Department. But that would be a different structure than the one we currently have. If that actually occurs, then we no longer need to appoint an advisory panel because there's a different process that is engaged at that point. So we are waiting to see how that works out. Now, that should all be resolved by the end of July, if not sooner. So my expectation, as soon as we know what the status of that is, is that we'll go out and appoint the advisory panel as soon as possible. But to answer your question, I don't know where things stand.

**COMMISSIONER DRISCOLL:** We're going to have to wait and see, and as soon as we do know something I would like to see some of the EMAB members be part of our advisory committee if that's the way it ends up legislatively.

**DR. SCHAEFER:** They might also be some other people who could serve on the panels.

**COMMISSIONER DRISCOLL:** Yes, absolutely.

**CHAIRMAN PEYSER:** Some of those outsiders that you're talking about.

**COMMISSIONER DRISCOLL:** Yes, absolutely. Very good people, as you know, on EMAB.

**DR. THERNSTROM:** My understanding is there's very little legislative interest in what the Governor had proposed.

**CHAIRMAN PEYSER:** In the House that seems to be the case.

**COMMISSIONER DRISCOLL:** We'll have to wait and see what the Senate says. We were told that they have to go out of business July 31 this year by rule, right, Rhoda? So I don't think you can go beyond July 31.

**DR. SCHAEFER:** Because this is an election year.

**CHAIRMAN PEYSER:** The rules can change. We are adjourned until tomorrow.