NATIONAL MATHEMATICS ADVISORY PANEL
STRENGTHENING MATH EDUCATION THROUGH RESEARCH

+ + + + +

MEETING

+ + + + +

Friday,
April 20, 2007

+ + + + +

Illinois Mathematics and Science Academy
1500 W. Sullivan Road
Aurora, Illinois 60506

+ + + + +

8:50 a.m.

PANEL MEMBERS:
DR. LARRY FAULKNER, CHAIR
DR. CAMILLA PERSSON BENBOW, VICE CHAIR
DR. DEBORAH LOEWENBERG BALL
DR. A. WADE BOYKIN
DR. DOUG CLEMENTS
DR. SUSAN EMBRETSON
DR. FRANCIS (SKIP) FENNELL
DR. BERT FRISTEDT
DR. DAVID C. GEARY
DR. RUSSELL M. GERSTEN
MS. NANCY ICHINAGA (NOT PRESENT)
DR. TOM LOVELESS
DR. LIPING MA
DR. VALERIE F. REYNA
DR. WILFRIED SCHMID
DR. ROBERT S. SIEGLER
DR. JAMES SIMONS (NOT PRESENT)
DR. SANDRA STOTSKY
MR. VERN WILLIAMS
DR. HUNG-HSI WU

EX OFFICIO MEMBERS:
DR. DAN BERCH
DR. JOAN FERRINI-MUNDY
MS. DIANE JONES
MR. RAY SIMON (NOT PRESENT)
DR. GROVER J. (RUSS) WHITEHURST

STAFF:
MS. TYRRELL FLAWN, EXECUTIVE DIRECTOR
MS. IDA EBLINGER KELLEY
MS. MARIAN BANFIELD
MS. JENNIFER GRABAN
MR. KENNETH THOMSON
MR. ROBERT GOMEZ

I N D E X

P-R-O-C-E-E-D-I-N-G-S

CHAIRMAN FAULKNER: I ask the person who's running the audio if it is true that I have to hold this button down while I'm speaking?

AUDIO TECHNICIAN: Yes.

CHAIRMAN FAULKNER: It is true, okay, and that means, I believe, that all of the speakers and everyone here at the panel will, when you do speak, have to hold the button down. Okay.

Okay, I'm Larry Faulkner. I'm chair of the National Math Panel. Vice-Chair, Camilla Benbow is next to me here. And we want to welcome everyone in the audience to this open session of the National Math Panel.

I'd like to begin by thanking the Illinois Math and Science Academy for hosting this open session. This is the sixth meeting of the National Math Panel. We are holding these meetings across the country in various locations, geographically distributed, but we've tried pretty consistently to associate the meetings of the Math Panel with institutions in locales that symbolize high achievement in the academic enterprise. The Illinois Math and Science Academy certainly does symbolize that.

Let me indicate that we have signing

services available. They are active right now. We can continue signing services through the entire meeting if there is a use for them. If no one is making use of them, we will discontinue. So I'd like to indicate whether this is a desire to continue signing services. Seeing no such indication, we will discontinue the services. If there is a need to institute them as the meeting goes along we can do that. Thank you.

Now, let me introduce Dr. Janice Krouse. Dr. Krouse currently serves as curriculum and assessment leader at IMSA. She is instructor of mathematical investigations II, III, IV, pre-calculus, and advanced-placement calculus. She has a Bachelor's Degree in secondary education in mathematics from Clarion University of Pennsylvania, a Master's in mathematical sciences from Clemson, and a Doctorate in mathematics education from the University of Pittsburgh. She is a member of the National and Illinois Council of Teachers of Mathematics, National Council of Supervisors of Mathematics. Dr. Janice Krouse will be representing IMSA.

DR. KROUSE: Thank you, Dr. Faulkner. Good morning. On behalf of the faculty and staff of the Illinois Mathematics and Science Academy, I welcome all of you here today.

I am honored to have this opportunity to greet such a distinguished group as you meet again to engage in this important work. I join you in recognizing the significant consequences of a quality mathematics education for the children of this country, as mathematics and critical thinking skills so profoundly affect their lives and their ability as responsible citizens to shape the human future. In my brief comments today I hope to share with you our vision of mathematics education and its power in shaping minds.

My colleagues and I take our role in influencing tomorrow's leaders very seriously. The quality of the engagement between teacher and student and between the student and the mathematics cannot be underestimated. It was for these reasons that the charter mathematics faculty and Presidential Awardees of the Illinois Mathematics and Science Academy invested their time, talents and energy into authorizing a pre-calculus curriculum named Mathematical Investigations for their students. With ongoing revisions and updates Mathematical Investigations, known affectionately as 'MI", is still taught here today. And I am proud to say that it is one of my primary duties to ensure the consistency and coherence of this curriculum and its delivery.

Charter math faculty and author of MI, Chuck Hamberg, often said, “If you stop when you get the answer to a problem, you miss half of the mathematics." It has been noted that one of the strengths of our program is the space we give students to solve a problem “85 different ways". It is that very notion of curiosity that drives learners to their full potential. It is our job as educators to believe in that potential and to create conditions in which it can be realized. What, then, is the role of the teacher in the MI classroom?

Largely, I am a guide. I imagine the impression of an observer to my MI classroom. It looks like a teacher's dream job. There is very little at-the-board lecture on some days and, instead, the observer sees me milling about the room, intermittently asking or answering questions of students who are working in small groups. Even first-time IMSA students sometimes wonder, when is she going to teach?

But next, I step back into my shoes as the teacher and the facade of a simple job shatters. Teaching is now far more exhausting than preparing lessons and lecturing. In the traditional format the teacher is almost always in control of what happens next. Everything is predictable, planned and polished. There is often a sense of, I taught it, so they now know it. Unfortunately, there is little way to actually validate that sense until a formal assessment is given, and by then, it's way too late for some students.

Mathematical Investigations (MI) invites learners into the science of mathematics through carefully crafted questions and problems. Students observe patterns and phenomena, make conjectures, test their hypotheses on new problems, and analyze their results. All the while students are engaged in conversations with peers and teachers about mathematics.

Ideas, probing questions, insights, and supporting arguments emerge daily. Through these conversations, students forge connections within and among mathematical concepts in ways that make sense to them. They utilize various forms of technology to explore and test their conjectures. Most importantly, they are not forced to merely absorb a neatly packaged explanation given by the teacher.

In fact, the word “teaching” takes on a whole new meaning in the MI classroom. It goes well beyond standing at the board and dispensing content, methodologies and algorithms organized in a manner that makes perfect sense to the well-educated and well-meaning teacher. It now means letting go, listening, assessing, reacting, questioning, probing, listening, reacting, clarifying, watching, listening, guiding, but not just telling, and again assessing every student in the room, every day.

There is a delicate balance of timing that must be maintained of when to let the students grapple with a new or difficult idea, and when to intervene, help them make necessary connections and to see the big picture. There is a constant need to think on your feet as students ask questions that even the seasoned teacher does not anticipate. There is a need for enough self-confidence and mathematical prowess to let the students watch you grapple with a challenging problem so that they can see you as a model problem solver, even if that means you make a mistake in front of them. This is something that the traditional teacher wouldn't dream of. There is a need to be able to answer students' questions with questions that lead them to the answers they thought they couldn't get. There is a need to hear the misconception that truly underlies the superficial I don't know how to do this one type of question.

Then, somewhere in the midst of the grappling and questioning, the synthesis begins. Students respond to the teacher's probing and challenging questions by refining their understandings of complex ideas. Ultimately, the forging of connections consummates in closure of sound mathematical ideas that students can transfer and apply to tomorrow's questions.

And once you think you've mastered all of that, you get a new class of students and you have the grand opportunity to start all over again. You find the balance again, perhaps in a different place, because all students are different, and teaching MI actually lets you see that and react to it. The MI teacher has the luxury of hearing students talk about the mathematics in their language, using their constructs. You learn to read how each student in your class thinks about mathematics, and you have the privilege of adjusting your instruction to suit all of those needs. That is simply impossible in a traditional classroom. Results on formal assessments are rarely surprises. Such tests are merely opportunities for students to demonstrate their knowledge in a more formal manner.

With Mathematical Investigations, we give our learners an opportunity to engage in the learning process in ways unlike any they have previously experienced. The mathematics environment here is clearly one of collaboration, making connections, and solving problems. These skills are absolutely fundamental to tomorrow's leaders.

These explorations are often aided by various forms of technology, including the TI-89 Titanium/CAS graphing calculator, Mathematica, Geometer's Sketchpad, Fathom, and the Internet. Technology enables students to actively pursue questions about mathematical constructs that otherwise would be unattainable. Further, today's students are engaged with technology so frequently, that to deny them this resource in their learning is asking them to divorce their natural environment from their schooling. Fluency with emerging technologies in problem solving will continue to be a critical, necessary and expected skill for our students.

Results? With over 850 students taking the Advanced Placement BC Calculus exam over the last seven years, we enjoy a collective average of over 4.6 on a five-point scale. Intel Finalists, Siemens winners, and inventors of Papal, Mosaic and YouTube are among our alum.

Certainly, we can lay claim to working with some remarkable students, but what we do in the mathematics classroom is applicable to a much wider audience. In 2003, IMSA mathematics faculty were called on as pedagogical experts to help a neighboring district to determine criteria by which a mathematics program would be selected for their high school that would invite a student-centered, problem-based, integrative and collaborative environment such as ours. After a complex and thorough process led this district to a selection, the process was repeated for finding an appropriate program for their honors students. After careful evaluation and critique, IMSA's Mathematical Investigations was chosen, and is now in its second year of implementation. The huge paradigm shift was not without its bumps in the road. But the benefits were evident to the teachers, even by the end of the first year.

One of their teachers stated: The most surprising thing that I encountered was how difficult it was to get students to try all the way through a problem. So many students waited for answers at the beginning of the year, and now they'd rather find out themselves than hear it from me. The challenges the students face make them think about their thinking. The students are willing to attempt any problem handed to them and understand many of the processes instead of just memorizing formulas.

Several years ago, an IMSA graduate recounted her experience as a first year physics student in a prestigious university's honors program.

The young, yet highly accomplished professor, who believed in the power of intimidation, began his lesson by asking the class if anyone remembered a very esoteric formula. As the IMSA graduate recalls, no one did, and much to the delight of her professor, the class nearly froze. Then, in the uncomfortable silence, she raised her hand and gave the formula. The professor, in amazement, asked her how she remembered it, and she said, “I didn't. In my school, we learned how to derive it.”

Ah, there is such power in giving students the space to solve a problem in a multitude of ways; in asking, “what if,” after the first answer is found so as not to miss half of the mathematics; of engaging students in deep, meaningful learning so that when the formulas fade, the understanding endures.

I want to thank Dr. Marshall, Dr. Faulkner, and the members of the Panel for this opportunity to speak with you today. I look forward to the recommendations that your research and wealth of experience bring to your final report.

CHAIRMAN FAULKNER: Thank you, Dr. Krouse. We appreciate the opportunity to be here and to hear from you today. Let's proceed now into the open session, which is going to be dedicated to public comment. I would like to begin by introducing the new members of the National Math Panel. There are three persons who have joined the Panel with this meeting here in the Chicago area.

Let me start by introducing Doug Clements, Professor of Early Childhood Mathematics and Computer Education at the University of Buffalo, State University of New York. Welcome Dr. Clements. Dr. Susan Embretson, Professor of Psychology at Georgia Institute of Technology, who is, where? Okay. Welcome, Dr. Embretson. And Dr. Bert Fristedt, Professor of Mathematics at the University of Minnesota. Welcome, Dr. Fristedt.

We have had comments from the public on an open basis consistently around the country. The comments that we have received have been done on a first come, first-served basis with the time available. We have found these comments to be quite useful, as we have received them over the period of the Math Panel's meetings during the last year approximately.

The nine speakers who will be speaking today were registered for public comments. The list is available to the Panel members in the notebooks under tab six. There is one person on the waiting list. Each speaker is limited to five minutes. There's a timer right there at the table where testimony will be made. Panelists will have the opportunity to ask questions of the speaker after the remarks are concluded.

Let me go ahead and open this testimony. The first presenter will be Henry Borenson.

DR. BORENSON: Mr. Chairman and members of the Panel, I thank you for this opportunity. My name is Dr. Henry Borenson, President of Borenson & Associates, Incorporated. Some twenty years ago, as a middle school math teacher, I was concerned with the difficulty students were having learning algebra abstractly. I determined to find a way to simplify the concepts and make them concrete and visual and to make them accessible to all grade school students.

After two years of experimentation working with children, including children with learning disabilities, I developed a system known as Hands On Equations. This is a system, which uses game pieces, such as those you see here, a flat laminated balance, and a specific sequence of ideas to enable students as early as the third grade to physically represent and solve algebraic linear equations. The types of equations, until then, were typically taught in the eighth or the ninth grade.

Since 1995, Borenson & Associates has conducted more than 1,500 Making Algebra Child's Play workshops throughout the United States. In these workshops, teachers of grades three to eight learn how to introduce the concept of a variable, the concept of an equation, the subtraction and addition property of equalities, and other key algebraic principles.

A key part of these workshops is a student demonstration with local fourth and fifth grade students. More than 1,500 times since 1995 the teachers attending our seminars have seen how, in three lessons, fourth and fifth grade students, including so-called low ability students, can learn to solve an algebraic linear equation such as $4x + 3 = 3x + 9$.

In a study to determine teachers' confidence level to teach algebraic linear equations to their lowest achieving students, Barbara N. Borenson (2006) discovered that only 17 percent of 751 teachers, from grades three to eight attending a Making Algebra Child's Play workshop, felt they would be successful using the traditional abstract teaching methods, while 98 percent expressed confidence of success if they were to use the Hands On Equations and materials. The study is shown in Appendix A of the handout.

In an ongoing series of studies involving multiple student characteristics and multi-site replications, supervised by Dr. Larry Barber, formerly director of research at Phi Delta Kappa, we have found significant pre to post-test gains for second grade gifted students, sixth grade regular students and ninth and tenth grade low achieving students.

Recently we completed a study involving four fifth grade inner city classes comprising a total of 111 students. The pre-test to post-test results showed a large and highly significant increase in scores. The combined mean increased in percentage terms from 44.8 percent on the pre-test to 85.3 percent on the post- test. On the three-week retention test, with no instruction in the interim, the mean was 78.6 percent. When compared with a pre-test score of 44.8 percent this increase was found to be statistically significant with a T-value of 13.71. We are talking about fifth grade inner city students succeeding with important algebraic concepts. This study will be found in the Appendix B.

We believe we have provided evidence that Hands On Equations learning system of instruction significantly and positively impacts teachers' self-confidence in their ability to introduce algebraic equations to their students. We have provided evidence that the program makes a measurable difference in student learning. We believe it is possible and it is important, as the previous speaker alluded to, for all students to gain the perception that mathematics is a subject that they can understand and a subject at which they can excel.

In Hands On Equations the students do not memorize a set of procedures in order to obtain an answer. They can use their creativity to apply general algebraic principles in the manner that best suits them. We ask the Panel to consider recommending Hands On Equations as a supplementary program that is effective in introducing grade school students to basic algebra. Thank you very much.

CHAIRMAN FAULKNER: Thank you, Dr. Borenson. Are there questions from the Panel? If not, thank you. I will now turn to the second --

DR. FENNELL: Mr. Chairman?

CHAIRMAN FAULKNER: Yes.

DR. FENNELL: Just one question. Dr. Borenson, the paper that you referenced --

CHAIRMAN FAULKNER: Push your button.

DR. FENNELL: I am. Thank you, Mr. Chairman. The paper that you referenced, will we have copies of that?

DR. BORENSON: Yes, in the handout. They will be available in the handout.

CHAIRMAN FAULKNER: Thank you. Let's go

now to the second presenter. It's Andy Isaacs. Is Andy Isaacs here? Apparently not. We will go to the third presenter, Cindy Jones. Is Cindy Jones here? All right, I will go to the fourth presenter, Patrick Thompson.

DR. THOMPSON: Does everyone have a copy of my testimony? Okay, there are some things in there. Mr. Chairman, Madam Vice-Chairman, Panel members, thank you very much for this opportunity to speak with you about the Panel's work. I will speak to five of the Panel's charges.

My first remark addresses charges one and seven; critical skills and skill progressions and research in support of math education. But it actually cuts across all of the charges that I listed, that I'm going to address.

The Panel has a significant task of responding to a list of charges that take skills as the primary component in mathematics learning when the notion of skill itself is hardly well-defined. Do you take skill to mean a child's ability to perform reliably a procedure when told to perform the procedure? Or do you take skill to mean a child's ability to have developed sufficient knowledge and appropriate flexibility of thought to solve most problems of a particular genre of problems, even those that might have subtle and nuance differences from any the students might have seen.

I am noting that I'm going to have to skip through some of what's in the prepared remarks because when I actually timed myself, looking in the eyes of people, I couldn't read as fast as when I was alone in my office.

Thus, it is incumbent upon the Panel to make clear where it stands with regard to what students should learn, and to justify that stance according to the pragmatic consequences that relative stances have regarding students' learning and preparation for future learning.

In regard to charges three and four, processes of learning and affective instructional practices, I offer an example from a current research project on affective models for secondary mathematics instruction.

We created an implementation of Algebra I in collaboration with one of the participating teachers in order to develop artifacts that would make concrete to the teachers what it was that we had in mind, that they had difficulty envisioning.

We also hope that these students would display proficiency in the algebra the teachers were accustomed to assessing, but display it as a consequence of understanding ideas well and not because of having memorized the prescribed procedure.

The students we taught were not in an honors program, thus they were taking Algebra I at ninth grade. Their computation skills were atrocious. They had no understanding of fractions. Their experience in mathematics was that teachers showed them procedures they were supposed to remember until the next test. Their feelings about mathematics were that it was a dehumanizing experience that no one in their right mind would choose to experience having had the option not to.

So our immediate question was what to do about their lack of skills given that our goal was to have them eventually engage with significantly mathematical ideas. Do we re-teach what they've already not learned? Well, we decided that we wouldn't, that we would move on. We began the year with no review and we designed the instruction by the seat of our pants, always guided by our goal of having them engage meaningfully with significant mathematical ideas and at the same time be able to pass their mandated tests.

We focused on central ideas prior to calculus curriculum like variation, covariation, rate of change and functional relationship. The appendices, by the appendices, I'm referring to files that are on the CD that I turned in. Those aren't printed in the materials that I gave you. The appendices contain examples of the kind of work we need to expect from the students.

Here, I'll give one example, to make a point. Actually, I'm running out of time so I'll let you read the example. But if you notice, it has to do with having students construct the sum of two functions, which are not defined by a formula, in their experience, but nevertheless, focuses on the idea that, in fact the sum of two functions is a function. Then we shared the definition and they became excited that they had dealt with such complicated functions and wanted a printout to take home to show their friends and parents.

Another point of what I say is that, in my opinion, this nation suffers not from a lack of research, but from a lack of imagination. It suffers from lack of imagination at all levels especially at the levels of policy and politics.

With regard to teacher training, at ASU our biggest problem is recruitment and retention. I give statistics about that in my testimony. One of them has to do with the fact that less than 30 percent of secondary math students who are required to take three semesters of calculus actually complete three semesters of calculus. In other words, we lose them. And we actually lose many of them from ASU, not just to other majors.

My time is up. I'm over-time. So I'll let you read the rest of my testimony.

CHAIRMAN FAULKNER: Thank you Dr. Thompson. Questions from the panel?

DR. BENBOW: Can you tell us what you actually did in the classroom to engage the students?

DR. THOMPSON: Can you be more specific?

DR. BENBOW: Well, you had them work the problems, but given that they didn't have the basic skills, how did you engage them in significant math without having had the basic skills already mastered?

DR. THOMPSON: Well, we focused on beginning with phenomenon, having them use literal symbols to represent phenomena. We focused on ideas of variable and variation so that variable stood for things that changed. The discussions were not about how to compute but how to represent. Computations flowed from that. Once you have a representation it's about how you would compute something. But the algebra that they wrote was algebra of representation, not necessarily algebra of computation, except when we looked at the mathematics of equivalence. Then we focused on the algebra of computation.

CHAIRMAN FAULKNER: Any other questions or comments? Thank you, Dr. Thompson. We now go to presenter number five, Kevin Killion.

MR. KILLION: Hi. I'm Kevin Killion. I hold a degree in mathematics. I have been a research VP in a marketing agency. I've written several commercial/statistical analysis products. And I operate a successful business in market and media analysis.

I became involved with math reform when I observed the difficulties my own son was having. Today I serve as director of the Illinois Loop, a 12-year-old organization of parents, teachers, school board members and others. Our Illinoisloop.org website is a valuable source about what is going on in schools and we have logged over 600,000 visitors.

First, I have a comment on standards. Calling one category of math programs standards-based is a ploy that tarnishes other programs as somehow being rudderless and adrift. I left over there my beloved American College dictionary. I looked it up. The word, standard, has 19 definitions. Similarly, there is no single standard for math.

Another weapon is to blame lousy math performance on attractable, dusty old methods.

Schools are constantly told to embrace change and teachers are exhorted to be agents of change. But the reality couldn't be more starkly different. Everything has already changed.

On our Illinois Loop website we provide extensive information about how math is taught in Illinois school districts, from Addison to Zion. This resource is well-used by parents in tracking what districts are doing. And here's what we found:

In Chicago, some 290 schools use constructivist math programs in early grades. On the flip side we've been able to identify only five, count them, five conventional Chicago public schools that use practice and mastery math programs. Plus there are another five schools that are charter schools offering Saxon Math.

With regards to the suburbs, the Illinois Loop has collected information on the math programs used in 118 suburban K-8 districts in five collar counties. We find that constructivist products form the math foundation in 77 percent of those districts. But even that only hints at what's going on.

On the north shore, or in Lake County or in some other areas, it's almost impossible to find any schools with anything but constructivist math. And across the area, the Chicago/Suburban area, we've identified only six districts out of 118 that make any use whatsoever of those math programs most recommended by practice and mastery reformers, such as Singapore math or Saxon Math. So much for the argument that parents in the suburbs already have the schools they want.

Now here's a twist. We've all heard of the dance of the lemons. Well, there's also the dance of the math lemons performed by districts unhappy with their math programs. As an example, District 39 up in Wilmette dumps Math Trailblazers and picks up Everyday Math even as District 109 in Deerfield drops Everyday Math to pick up and have a chance on Math Trailblazers.

Like Lois Lane who couldn't see the truth staring her in the face, these districts stick with constructivist math and merely substitute one program for another. We're sure not seeing any agents for change there. These districts are firmly mired down with a philosophy that they refuse to abandon.

In the course of our work at the Illinois Loop we receive hundreds of messages from parents. Many of them are concerned about constructivist math programs in their schools and what these programs are doing to their kids. I'll close by reading just a few snips of what parents are saying. I implore you to listen to the passion and the concern expressed.

A Glencoe mom tells us that math problems here are bad and getting worse. A Skokie math teacher told us that this series has been a dismal failure in teaching math. A Homewood parent tells us that the math program there is the most confusing, ridiculous method she's ever seen. She couldn't believe parents are accepting this and how sad it is for their children. A Glenview couple writes that the math program there stinks. A Downers Grove parent wrote to us, "It is beyond belief that so many parents can be so upset at the situation and yet be paralyzed."

A Hinsdale parent told us that more than 40 percent of parents pay tutors up to $50 a hour to teach their kids properly. A Naperville mom fears that when her daughter finishes in this school system, she will be well experienced in arts and crafts, but she will lack the ability to make change. A parent laments that as the result of the math problem in her Lake Forest school, "You can't get your kid into Kumon classes around here. When will they learn?" By the way, in Naperville there are nine Kumon centers in the area. A Crystal Lake parent wrote to us, "Everyone I have talked to thinks this program is horrible and their kids are struggling." A Batavia couple says, "This trend needs to be stopped now before we have a complete train wreck.” A Plainfield parent says, “I think it's the most absurd form of education that I've ever seen.” And a Yorkville mom sums it all up by saying, “Help, how can I save our children from this blight?”

CHAIRMAN FAULKNER: Your time has just expired.

MR. KILLION: I just did. A Yorkville mom says, “Help, how can I save our children from this blight?” Members of the Math Panel, thank you for your concern.

CHAIRMAN FAULKNER: Thank you, Mr. Killion. Questions or comments? Yes, we have one here.

DR. SIEGLER: I grew up in this area so I know most of the suburbs that you mentioned are quite affluent suburbs and the parents aren't usually shy about organizing if they have a strong opinion.

If these are representative of parental views, what do you think is keeping school board members from being elected who want to change the current system?

MR. KILLION: I don't think there's sufficient time to go into the problems of school board elections here. Suffice to say that these are real opinions representative of hundreds that we get at the Illinois Loop. They are suffering with what's going on with their kids. If somebody else believes in a different way of doing things and they want to choose a program for their kids, that's fine, but these are parents who are suffering.

CHAIRMAN FAULKNER: Are there any other comments or questions? Let me go ahead and proceed to the next presenter, Jack Rotman.

MR. ROTMAN: Let's see if I can master the technology, is that okay? Is the microphone working? No? Is that better, okay, thank you.

To briefly introduce myself, I am Jack Rotman. I have been a professor at Lansing Community College in Michigan for 34 years. I have been active in American Mathematical Association of Two-Year Colleges (AMATYC). I currently chair the developmental mathematics committee of that group. And I was a contributing writer for the 2006 standards document, Beyond Crossroads.

I have three questions for the panel, which are the basis for my remarks. One: Are sufficient and necessary conditions present in the schools to provide mathematics learning for all students? Two: Are there barriers outside of the education system that substantially limit the learning of mathematics for some groups of students? Three: Do we plan for the system, which provides a second chance for students who did not learn sufficient mathematics in the schools?

On the first question, are there sufficient and necessary conditions present? At the most basic level, students must stay present and attending in order to benefit from the curriculum. At the secondary level we are all aware of the substantial problem with drop-outs. However, there are also a lot of absences in the schools. Studies show that seven percent of the students were absent on a given day and that was only for unexcused absences. For students who are present we need to be concerned about how much they are actually attending. An optimistic study estimated that students were in attendance and with material 65 to 75 percent of the time.

In a different study of various methods of teaching, the only method that increased student attention was the debate/discussion method. The group learning methods only increased attention a little bit.

On the second question, are there barriers outside of schools that limit opportunities? The Panel has discussed the concept of stereotype threat, which is one of those barriers. I would encourage the

Panel to consider broader viewpoints of these issues. One of these viewpoints is called critical race theory.

Critical race theory begins with the assumption that racism is embedded within the social structure and analyzes information from that viewpoint. Critical race theory suggests that the achievement gap that we talked about is really an opportunity gap. A more radical view sees standardized testing as a means to justify differences.

Also, some researchers have documented a default trajectory towards dropping out in certain types of communities. In addition, in some regions we again have schools that are separate but not equal due to policies such as schools of choice and other issues. This segregation results in a situation where the Lansing high schools are 70 percent minority, while the Lansing area itself is only 35 percent minorities.

I would also encourage the Panel to consider other barriers that exist outside the education system. For example, mathematics still faces the barrier that it is acceptable or even desirable to be "bad at math." Will we hear the President say that qualitative, quantitative literacy is a personal value for me? How about our role models in entertainments and sports? Are they going to say my mathematical skills allowed me to accomplish what I needed? Or will you see Ben Wallace helping middle school students with their mathematics? Or do we see these pupils say, math was always hard for me too?

On the third question, the back up system, the second chance. Most countries don't have our community college system. Even the community colleges offer a second chance for many adults to learn the mathematics they need. However, the country hardly has a systematic plan for this approach. Outside of the work of American Mathematical Association of Two-Year Colleges (AMATYC) and a little bit of The Mathematical Association’s work, nothing systematic is done beyond the state level.

I will suggest the Panel consider community colleges as part of the system and that we be included in the dialogue. We provide a recruiting ground for mathematics and science fields. Also, we offer a response time measured in one to three years instead of 12 years for K-12 schools. I would think our involvement would be appropriate.

As we consider our work to strengthen mathematics education I hope we can establish those minimal conditions for learning, look at barriers to learning outside of the schools and include community colleges in our discussions. Thank you for your attention and the opportunity to address the Panel.

CHAIRMAN FAULKNER: Thank you, Professor Rotman. Questions or comments from the Panel? None. The seventh presenter is Ken Indeck.

MR. INDECK: My name is Ken Indeck. I am a high school math teacher with nearly three decades of experience and I'm speaking as a representative for the Illinois Association for Gifted Children. My remarks are primarily anecdotal because it is important for me to communicate the realities as viewed from within the school system. I am confident that similarities exist in most educational settings.

One of the hallmarks of gifted education is the notion that one size does not fit all. In Illinois the same content benchmarks are used to assess all students. For the bottom third of the academic spectrum these benchmarks are a stretch, often unrealistically so. For the top third, these students have often surpassed them.

Last year I was talking through some curricular improvements we could implement for bright students in our building. Before I finished, the administrator I was speaking with, stopped me and said, you're not going to want to hear this, but that's not going to help us meet AYP, those kids will be fine. We need to focus on raising the scores of the students who will help us. Unless you think otherwise, that administrator is an excellent educator.

As a parent I was thrilled when our son's third grade math teacher told us how proud she was the entire class had completed both the third and fourth grade material. Imagine my shock when we found his fourth grade math teacher was teaching the fourth grade curriculum, knowing full well the students had been through and mastered that content, simply because she was not able to teach the fifth grade material. Half that class lost interest in math. By sixth grade there were a handful of students who were still excited about math and ready for algebra, but they were not allowed to take the course because the junior high didn't offer it. My son is now in tenth grade and I say with mixed emotions, he is doing fine.

I envision four entwined approaches to improving our current state of affairs in math education. First, advocate for the use of best practices. Acceleration is important but it is not enough and absent coordinated sequence spanning years it can even be detrimental.

Few high school math teachers are

knowledgeable about differentiated instruction and fewer still are skilled in its implementation. For many high schools the gifted curriculum is synonymous with AP course offerings. Well, this might be a starting point. We know better. Yes, we want our brightest and most able students exposed to age-advanced concepts. However, those students thrive when they are also in a rich environment that helps them see connections to other topics in the curriculum, and where they are allowed to explore how those connections can be put to use making the world better by improving people's lives.

Second, encourage and support the educators who take reasonable professional risks. The current practice of looking for significant improvement over short stretches of time does not realistically encourage a teacher to switch from one set of techniques to another, even if the new set is extremely promising, when it will likely take on the order of five years to master those skills, and another five to ten years to become expert in their use.

Third, it is essential to provide significant support for research. In education we need research regarding instructional practices. We need to know more about how grouping students and sequencing topics influence learning. It is important to develop broader assessment practices, practices that extend beyond recalling facts and solving one or two step problems. In order to maintain our nation's leadership in the areas of science and technology it is essential to support math, the research in mathematics, science and their applied fields both through academia and industry.

Finally, it is crucial that we do a better job educating the public about the educational enterprise as a whole. The typical adult non-educator does not fully understand how hard the work is and unlikely has an realistic set of expectations both for what our schools can provide and how the educational growth of students can be documented.

We have before us opportunities for establishing long-term leadership for the economic strength and for improving the quality of life for our nation and the world. That leadership is likely to come from students at the top end of the academic spectrum, who are well grounded in math and science and who recognize the connections between those subjects and the broader world around them. The notion that we are doing fine is not good enough.

Strengthening the educational system should prompt increased achievement for all. Closing the achievement gap should not translate to holding hostage the education of our most able students. If we compare students' performance to their own capabilities, as the mission statements for most schools suggest, it is the bright students who fall short and are furthest from reaching their potential. We must do better. Please help us. Thank you.

CHAIRMAN FAULKNER: Thank you, Mr. Indeck. We appreciate your comments. Comments or questions from the Panel? Wilfried?

DR. SCHMID: You introduced yourself as a teacher of gifted children. At what kind of school do you teach?

MR. INDECK: I teach at a regular high school. I am not teaching gifted classes at this point. I was at one time in my career the curriculum and staff development coordinator for High School District 214 for their Talent Development Program.

DR. SCHMID: So at that point you were designated as a teacher of gifted children?

MR. INDECK: In our district we don't call gifted children. It is a program for developing talent.

DR. SCHMID: I see.

MR. INDECK: That's the closest we have in our district.

DR. SCHMID: In any case, but you were designated as such?

MR. INDECK: Yes.

DR. SCHMID: And no longer are? Did the policy change?

MR. INDECK: Yes, I don't currently hold that position.

DR. SCHMID: Does anybody else?

MR. INDECK: No, the position was disbanded because it doesn't help them meet AYP.

CHAIRMAN FAULKNER: Valerie?

DR. REYNA: Thank you. What do you think the barriers are to really having two goals in mind at the same time, the adequacy goal and excellence as a goal as well? You make the argument yourself in your own testimony that these two are not exclusive. If we're focusing on one of the goals, why does that mean the exclusion of the other? What do you think the barrier is there?

MR. INDECK: There are multiple barriers, but it seems to me that when we're in a system that tries to get all students to a certain level and doesn't look for growth on the part of all students, that once students are to that particular level, there's very little incentive within the structure itself to move those students forward. The focus is more on moving those students who haven't reached that benchmark, to the benchmark.

CHAIRMAN FAULKNER: Any other questions or comments? Thank you, Mr. Indeck. We now turn to presenter number eight, Sarah Delano Moore.

DR. DELANO MOORE: Good morning. My name is Sara Delano Moore and I'm the Director of Mathematics and Science at ETA/Cuisenaire. ETA/Cuisenaire is a leading publisher of supplemental instructional resources for mathematics, science and literacy. For over 40 years our company has pioneered the development and effective use of hands-on materials or manipulatives to improve student learning outcomes.

I am here this morning to share my thoughts on the role of manipulative-based instruction in mathematics, and I will begin by briefly sharing my own background.

I am a fourth generation teacher, although the first to teach mathematics. My undergraduate education focused on molecular biology, so I am a scientist by training. I taught mathematics and science in middle grade schools and have worked in higher education as well teaching both mathematics methods courses and curriculum. My research in writing has focused on the use of award winning and high quality literature, alongside hands-on experiences, to teach rich mathematics and science at all levels.

ETA's products and associated professional development training have always been grounded in the belief that children learn mathematics by doing it in active, hands-on ways. We are fortunate that this belief has a long research base to support it.

The three part learning cycle we use to discuss instruction with manipulatives includes phases called concrete, representational and abstract. Jerome Burner's work talked about a similar cycle as inactive, iconic and symbolic. Most recently Michael Batista used the terms, action, reflection, and abstraction. In all cases the basic idea is that children must first have hands-on experiences with the math and then use the representational phase as a transition to the abstract more formal mathematics.

There is no question that children need to be computationally fluent. These children must also understand the mathematics behind the computational procedures they use. I love mathematics. I earned good grades in math class at school. I'm not sure, however, that I genuinely understood mathematics until I learned to use manipulatives to teach math.

I had my first “ah-hah" experience in

mathematics in my early 20's when I learned to multiply two digit numbers with base ten blocks. I finally knew what was really going on when I wrote down all those numbers years ago in fourth grade. I saw the connection between multiplication of whole numbers and binomial multiplication in algebra. Math became a connected whole for me.

We don't know what problems our students will need to solve as adults. We can be certain they will need problem solving skills. They will also need the confidence they can solve problems successfully. Children learn by making connections between the familiar and the unfamiliar. Our role as teachers is to guide children toward the connections we want them to make.

Manipulatives provide a bridge between the concrete world of a child and the abstract concepts of mathematics. They may also serve as an enticement to learn math, which does not, on the surface, appear engaging. By using the manipulatives, literature, and other active instructional resources, children can be drawn into the world of math and find success there. Every child must find meaningful success in mathematics, and we must use every resource we have to ensure this happens.

Effective use of manipulatives is one

resource to help children find success in mathematics. Children must conduct structured investigations and work towards an understanding of procedures and strategies that can be generalized.

Unfortunately, manipulatives are too often used as hands-on worksheets with teachers telling students exactly which piece to touch and where to place it as they act out the traditional algorithm. Professional development is critical if teachers are to use manipulatives as the powerful tool research shows them to be.

For all of us as teachers it is a great day when a student has an "ah-hah" moment of learning. One of the joys of my job working in professional development is to see that same "ah-hah" from adults as they see mathematics taught with manipulatives and understand, sometimes for the first time, what was really going on back in elementary school when they obediently memorized the sequence of squiggles on the page which represented a mathematical idea.

Manipulatives are one of the most powerful tools in a teacher's arsenal for helping students learn mathematics well. I urge the Panel to ensure these "ah-hah" moments continue in classrooms by supporting the use of manipulatives in mathematics instruction in their report. Thank you.

CHAIRMAN FAULKNER: Thank you, Dr. Moore. Are there questions? There obviously are. Tom?

DR. LOVELESS: You mentioned the research that supports manipulatives? Could you describe a couple of, or even just one piece of research that you are convinced is persuasive in that regard?

DR. DELANO MOORE: I made an intentional choice today not to do the foot-noted presentation and to talk instead. There are a number of pieces of research, for example, on the use of base ten blocks and various models. Also there is the work that Karen Fuson and her colleagues in John Bransford's group about how students learn text. For example, they talk about the role of working from a concrete model even if it’s a sketch.

And as I say to teachers who say they can't use manipulatives on most state tests (Wisconsin I think, is an exception, but there aren't many), “No state has banned scratch paper.” So when they learn to make those sketches, they can then, as our opening speaker said, derive the formulas. I'd be happy to provide additional, more formal work to you if you'd like.

DR. LOVELESS: Just one follow up, would you agree that the goal would be for students not to have to depend on manipulatives eventually?

DR. DELANO MOORE: In my experience, as children learn the math, the use of manipulatives really self-extinguishes. They reach a point where they can use most often what are common algorithms, perhaps an algorithm of their own adaptation, but can do that work independently. The manipulatives serve as a tool to bridge between their concrete world and concrete thinking and the more formal mathematics that they will need in life.

CHAIRMAN FAULKNER: Other questions or comments? Valerie?

DR. REYNA: Are you familiar with the research of David Uttal on the use of manipulatives, U-t-t-a-l?

DR. DELANO MOORE: I don't believe I am.

DR. REYNA: And by the way, I hate to put you on the spot on this.

DR. DELANO MOORE: That's quite all right. Saying “I don't know” is an okay thing to do.

DR. REYNA: It certainly is. If you do take a look at the research and want to communicate with the Panel about the research, there are ways to do that. I would be interested in your reaction to that. It may only be a question of at what age manipulatives are appropriate to use.

DR. DELANO MOORE: All right, I will take a look.

CHAIRMAN FAULKNER: Anything else? All right, thank you, Dr. Moore. Let me indicate the procedure that we'll follow. Number nine, Barbara Wilmot, is next. Let me ask her to come forward. She was the last signed up member. Since we had some who did not arrive I want to indicate that we will next take Janie Zimmer who was the person on the waiting list. And we will then proceed to pick up number three, who has arrived, and that's Cindy Jones. So we will go to Barbara Wilmot next.

DR. WILMOT: Thank you, good morning. My name is Dr. Barbara Wilmot. I've worked in mathematics education from the elementary to the university level for 45 years now. I taught at Illinois State and directed a state professional development program there. Now I'm an independent consultant and administrator for a grant that supports and monitors central Illinois schools that don't make AYP year after year.

I've worked with over 100 districts and stopped counting when I'd given 1,200 professional development workshops in almost every state. I'm speaking this morning for myself and for Learning Resources, which is a leading provider of hands-on classroom materials. I often use their materials in my professional development sessions and have partnered with them to create this mathematics manipulatives handbook, which they hand out free by the thousands of copies in order to help teachers.

Today I speak really on behalf of millions of students with language barriers or special needs, many of whom are in mainstream classes. And yet for the most part No Child Left Behind holds these students to the same level of expectation as other students. How can we level the playing field for them in learning mathematics?

I'd like to share three points supporting the fact that hands-on learning tools and related professional development help English language learners and students of special needs deepen their understanding of mathematics and increase achievement.

The first point is similar to hers that manipulatives allow students to build, model and create multiple representations of mathematical concepts and, therefore, help them meet benchmarks. Whether we use NCTM or state standards as a guideline, “build,” “model” and “create” are verbs that appear at almost every grade level.

Other verbs such as “describe,” “verify” and “generalize” also happen if engaging tasks are offered for students. Certainly it is difficult to meet these outcomes without using manipulatives.

Meeting benchmarks and developing a deep understanding require that students explore multiple representations of mathematical concepts. Students aren't likely to fail if they only learn fraction concepts, which is a pre-requisite to learning algebra, in one representational format. Just imagine if learning fractions meant only drawings of pizza slices, and unfortunately that's the reality in many classrooms.

But in schools like West School in Glencoe, Illinois, teachers, like math specialist, Laura Menonski are using multiple formats. Laura recalls modeling the concept of two thirds to her students, and she could tell by the glazed look on their faces that her explanations, instead of drawings of set models, weren't enough. Then she brought out manipulatives like the fraction spheres and tower tubes to show two thirds in multiple dimensions. And when students modeled and saw the different formats they literally said, oh, and explained to her what she was trying to tell them all along.

My second point is that manipulatives allow students with limited language abilities and/or special needs to understand simple and complex mathematical concepts and to actually demonstrate their knowledge. Manipulatives enable English language learners and students with special needs to see concepts being modeled even when the students are unable to understand the teacher's words.

Physical models also allow for assessment. Students can build the representation and demonstrate knowledge of ideas when they aren't ready to communicate via symbols or words.

Chris Triola, a sixth grade teacher from General McLean School District in Edinboro, Pennsylvania says manipulatives allow his students with special needs to develop “insights and connections not available through paper, pencil or lecture.”

My third and final point is that high quality professional development is absolutely essential to learn how to integrate manipulatives and a variety of strategies and techniques into the curriculum to differentiate the instruction for each student. Teachers believe, in general, manipulatives are highly effective, yet few actually use them and fewer yet know how to use them correctly.

Manipulatives are most effective when the students use them to probe and make conjectures and generalize about a mathematical problem.

CHAIRMAN FAULKNER: Please wrap up, your

time's expired.

DR. WILMOT: Oh, I'm sorry. I have found that at least 100 hours of professional development are necessary to make teachers comfortable with this. So in order to meet the various needs, learning styles and abilities, I hope that you'll really think about the manipulatives as well as the professional development piece in your recommendations. Thank you.

CHAIRMAN FAULKNER: Thank you, Dr. Wilmot. Questions or comments from the Panel? Diane?

MS. JONES: I have a question. You obviously have a lot of experience in teacher professional development, and you mentioned the term high quality. You know, the U.S. government spends millions of dollars every year on teacher professional development and yet it's very hard to distinguish high quality from low quality. Could you give us some guidance? In what way is professional development best delivered and how should we be assessing federally supported teacher professional development to distinguish high quality from moderate or low quality opportunities?

DR. WILMOT: Wow, that's a good question. First of all, I really think that we’ve given up on the one-shot professional development. However, it’s useful for awareness and for disseminating information for an introduction to something.

But I really think that it has to be district based and/or school based. It has to be long term. There has to be support and an administrator. The best situation that I've ever had is when the administrator is there for every class or session that I have and then goes in and says, okay, would Tuesday or Wednesday be better for me to see how you're using this. So I think that the use of it is really good. And I think the keeping of data, both on student achievement and teacher opinion. Having teachers journal and reflect is a vital part of it too. So there are just a lot of phases. But just the coming in and going out doesn't help, you know. Less than ten percent of change actually happens in the classroom with that.

CHAIRMAN FAULKNER: Other questions or comments? Thank you. We'll go to Janie Zimmer.

MS. ZIMMER: I'm glad I'm not too technology challenged. This is interesting. Thank you for the opportunity to speak to you today. I am Janie Zimmer from Research Based Education speaking on behalf of National Council of Supervisors of Mathematics (NCSM). I serve on their board.

This morning I would like to discuss an issue that is critical in mathematics education. The critical issue is equity, the opportunity for and the expectation that every child will be successful in mathematics and will have the opportunity to reach high levels of mathematical content.

Schools and teachers do have that expectation for a lot of our children. And we think that we have this expectation for all children when we profess to permit children into higher levels of math classes beginning with Algebra I, if they are prepared and ready for that rigorous work. We profess we want every child to be successful, that is to get good grades.

In the meantime, we continue to sort and select which students will go into which high level classes and which students will go into the low level or remedial Algebra I A/B classes. In many schools educators create classes into which they place students according to their performance on state assessments. Or they create inclusion classes that contain both general education and special education students, frequently without support. But does that act in itself create equity?

In the words of a Pennsylvania teacher, I expect very different things from the lower level or inclusion class than I do from other classes. Individual Education Plans (IEP’s) send the message that a student does not have to perform in the same way as my other students. Isn't that holding a different expectation? What I am communicating is that some of my students are not smart enough to do the same high-level work. Yet how are students who enter the ninth grade with fourth grade mathematics skills able to do the ninth grade high algebra content? How are they able to do the same high-level work of on grade level students who are entering that same algebra class?

A school district of about 50,000 students in Maryland has grappled with this issue. Today, all students in the middle school are placed in on-grade level classes with added support for struggling students. In all 12 of their high schools, all incoming students take Algebra I as the minimum class. Students with IEP's or 504 plans are included in these regular classes.

In addition, high schools provide an extra support seminar as part of the schedule of students who need extra help. These classes are assigned two teachers; a math certified teacher and a special education teacher. The classes have a student/teacher ratio of 10:1, and they are co-taught by both teachers. An observer walking in would most likely not be able to tell which is the special education teacher and which is the general education math teacher.

They have had much success with this program. All 12 high schools have achieved AYP in mathematics for all populations. Overall in the district, the special education students of the extra seminar class had a pass rate on the state algebra data analysis assessment that was 17 percent higher than the general population for those algebra classes. That is, the group of the special education students actually outperformed the general population.

In addition, special education students who were in the extra seminar class had a pass rate that greatly exceeded the pass rate of peer special education students who had not been placed in the extra seminar class. They exceeded by 36 percent in one school and by 33, 27, 25, and 21 percent in similar schools.

As we look throughout the country we see other successful programs regarding equity in place. Most special education students are not intellectually challenged but they are challenged in many other ways. Equity is on the plate of most mathematics educators yet they need to grow and expand their understanding of the deep implications of this principle.

We realize that equity in itself is not

the mission of this Panel. But we ask you to take to heart our information and address equity in every facet of your work. Address the equity not only for students with special needs, but also for students who are speakers of other languages, who are economically challenged, who have families unable to provide support, who seem unmotivated or who, in some other way, do not fit the norm. NCMS used to consider this and we invite you to call upon us to inform your work and provide support in any way that we can.

CHAIRMAN FAULKNER: Thank you, Ms. Zimmer. Questions? Right here, Vern?

MR. WILLIAMS: You said that all of the students took algebra in ninth grade, but did some of the students take algebra in eight grade and then geometry in ninth grade?

MS. ZIMMER: That's correct. The school system in question is Howard County Public Schools and they do have a gifted program in place where a lot of the students, or a number of the students in seventh and eighth grade actually take algebra and geometry. They may come into ninth grade taking geometry or they may come into ninth grade taking Algebra II.

MR. WILLIAMS: Okay, so they have basically sorted the population starting in seventh and eighth grade?

MS. ZIMMER: Yes, they have.

MR. WILLIAMS: And my other question is, the test that they used as a comparison, was it the Maryland State Algebra Test?

MS. ZIMMER: The state test in Maryland is an Algebra/Data Analysis test and that is the test that they used.

CHAIRMAN FAULKNER: Diane?

MS. JONES: I'm quite familiar with Howard County and the growth of the number of Huntington and Sylvan Learning Centers, tutoring centers that have grown in Howard County in the past five to ten years. Was there any collection of data in this study in terms of the number of students involved in this study who were also receiving supplemental tutoring by the many Huntington and Sylvan centers that now exist in Howard County?

MS. ZIMMER: I'm not aware that there was that correlation made.

CHAIRMAN FAULKNER: Valerie?

DR. REYNA: Are the data that you just presented here going to be made available to the Panel?

MS. ZIMMER: I do not have this data in my possession right at this time, but I can get them and send the reports to the Panel.

CHAIRMAN FAULKNER: Anyone else?

MS. ZIMMER: If I could just add one other thing. The co-taught classes were classes where there was a lot of professional development for the teachers. So the special education teachers were brought up to speed on the content in mathematics, which we find to be a problem across the nation.

CHAIRMAN FAULKNER: Thank you. Okay, we are now going back to pick up number three, Cindy Jones.

MS. JONES: I come to you from Providence, Rhode Island where I am a curriculum coordinator for mathematics. I work in a largely urban community with a large immigrant and Latino population. My purpose for coming here is just to describe some aspects of the professional development that we've engaged in as teachers that I feel is very effective.

Since the beginning of my teaching career I've always had a love for data. This interest started in 1998 when, in my first year of teaching, my principal informed me that a RIDE, Rhode Island Department of Ed official was coming to observe my class. The Rhode Island official that came to observe me did not revoke my teaching certificate. Instead, she invited me to join her workshop.

The next three years, working with the

Rhode Island Department of Ed Office of Assessment Accountability Teacher Committee, I learned so much. I became sold on the idea of using rubrics to assess student's work. I was also sold on the idea that our assessments and what we teach should be closely aligned to state standards. I became proficient at looking at standardized test results to help form my instruction.

The SIP model, the Standards in Practice, which is part of my appendices, has become an essential piece of professional development for teachers, administrators and curriculum coordinators. The SIP model encourages colleagues to come together and discuss student work in terms of how the work demonstrates proficiency, the math concepts or grade level expectation and the Rhode Island standards being targeted.

Colleagues are prohibited from discussing the student, but rather discussing the work itself. In the SIP model, at first everyone assesses a bunch of student work on his or her own. Then in small groups, colleagues have discussions regarding the grades they have assigned to each piece of work. When discrepancies arise, colleagues are asked to reexamine the student work and the rubric to come to an agreement. The process allows educators to share ideas and their perspective with one another.

A typical rubric, I'm sure you're familiar, is usually a one through four. One is below proficiency, two is partially proficient, three is proficient, and four is proficient with distinction. As you can see the use of rubrics has permeated every aspect of our school community. It has been a powerful tool for us as teachers to keep the main thing, the main thing. And more and more we are learning not to judge student work based on personal biases or family history, but more on what the student was actually able to produce.

Since then I've become a math coach. Being a math coach allows me to integrate standards and assessment into my practice. One of the things I do often in team meetings is look at the New England common assessment programs release items, which are released by the Rhode Island Department of Education annually. Twenty-five percent of that exam is released annually.

And one of the things that we do with these release items is we align them to specific grade level distinctions and Norman Webb's depth of knowledge levels. Then we compare what we have to the release test answer page.

Norman Webb's depth of knowledge of

mathematics consists of four levels of proficiency. The fourth, which is level four, is the most rigorous type of assessment item. It requires more high order thinking skills than the other three.

The New England Common Assessment Program, otherwise known as NECAP, does not assess at level four. The first depth of knowledge level, assessment items may consist of simple recall or recognition of facts or math terms and application of a well-known algorithm. The other levels require more and more thinking skills, such as comparing/contrasting. Depth of knowledge two is more of the comparing/contrasting. Justifying and making conjectures is depth of knowledge level three. You'll find reference to these different levels in my appendices.

Integrating depth of knowledge into assessment items makes room for rigorous instruction. As a result teachers have to go beyond just hitting the surface of math concepts. We have to build the kind of understanding that allows kids to make conjectures and draw conclusions. As a result we know we have to spend more time on math concepts and we have to introduce them in many different contexts.

I would love to see more of my colleagues and myself receive professional development in the ways I've described above. I believe it has helped raise our school and district student achievement scores in mathematics and empowers us as teachers to own what we teach.

CHAIRMAN FAULKNER: Thank you, Ms. Jones. Questions or comments? There are none, thank you.

We will reassemble at I think 10:25 for the session which will involve the reporting of task groups. We are now concluding the task group.

(Whereupon, the above-entitled matter briefly went off the record.)

CHAIRMAN FAULKNER: This letter has been put in my hands and I want to convey it to you. This is a letter from the Vice-President Eric McLarin at IMSA.

A proclamation was issued by the Governor asking for all citizens of Illinois to join in a moment of mourning and ring bells in memory of those who lost their lives earlier this week at Virginia Tech. The moment will be observed today at 11:00 a.m. Dr. Gebble, 1980 Virginia Tech graduate with a Ph.D. in microbiology will lead the IMSA community via the public address system. The Governor's proclamation is printed below. I'll read the proclamation.

Whereas, the Commonwealth of Virginia and the United States of America suffered a great tragedy on April 16, 2007, when 32 people were murdered and dozens more were injured on the campus of Virginia Tech in Blacksburg, Virginia. And whereas the State of Illinois grieves with those who lost loved ones on that day. And we pray that they and the entire Virginia Tech community can someday find peace and solace in the wake of this senseless act of violence. And whereas in the words of Virginia Governor, Timothy M. Kaine, “April 16, 2007 will be remembered in the hearts and minds of Virginians and all Americans for the rest of their lives". Indeed this is a tragedy that our nation will never forget and we come together as a people to mourn with the victims' families. And whereas Governor Kaine will declare a day of mourning in Virginia on April 20 highlighted by a bell-ringing ceremony at noon Eastern time in honor of the victims of the Virginia Tech tragedy and whereas Illinois is humbled, yet saddened, to join in this solemn observance and will hold a bell-ringing ceremony in accordance with Governor Kaine's declaration. Therefore, I, Rod R. Blagojevich, Governor of the State of Illinois, do hereby proclaim April 20, 2007 as a day of mourning for the Virginia Tech victims in Illinois, and I encourage all citizens to join in the ringing of bells at 11:00 a.m. Central time in memory of those who have lost their lives on that dreadful day.

That will occur at 11:00 o'clock and we will stop what we're doing and simply be a part of it.

With that let me turn the program over to Vice Chair Camilla Benbow who will preside in this next section.

VICE CHAIRPERSON BENBOW: We now move to the open session to hear progress reports from the various task groups. For those of you who may not have been following the National Math Panel's work too closely, let me just give you a little bit of background how we are conducting our work.

The Presidential charge asked us to address several questions, and we decided that the best way to organize our work and be most effective would be to form first, four task groups to address the questions in the presidential charge.

The first task group is the Conceptual Knowledge and Skills Task Group. The second task group is Learning Processes. The third task group is Instructional Practices. The fourth one is Teachers. Those began right away.

It was always the intent that we would have an Assessment Task Group as well, but we wanted to make some progress on the first four before we formed the Assessment Task Group. The Assessment Task Group was actually formed at this meeting and has had already some meetings. And today, now, we will report out what is the progress of their work so far.

So I am going to ask each task group to come forward, either the Chair or several individuals in the task group are going to give a report of our work so far. Approximately a third of our work has been looked at. I'm not sure it's exactly a third. And we hope to continue reporting out bits and pieces at the next meeting and again in St. Louis.

So the first task group that I ask to come forward is Conceptual Knowledge and Skills and the chair of that task group is Skip Fennell.

DR. FENNELL: Good morning. I'd like to acknowledge my task group and some others who have contributed to our work along the way; particularly task group members Dr. Sandra Stotsky, Dr. Larry Faulkner, Dr. Wilfried Schmid, and Dr. Liping Ma. Then we have other members of the Panel who have assisted in assembling our report to date including particularly Dr. Hung-Hsi Wu.

So we are essentially addressing three questions, the first one being what are the major topics of school-based algebra as we know it. Our analysis includes a review of states with standards for Algebra I and Algebra II courses, the relatively recent grade 12 NAEP objectives, the two related initiatives from Achieve, the American Diploma Project benchmarks, as well as their end of course test in Algebra II, and Singapore Mathematics Curriculum for grades seven through ten.

I'm hesitating here because I'm noticing that several members of the Panel are getting cups of coffee and one of you better grab one for me. Okay, I lost my train, sorry about that.

We're also looking at additional international comparisons and major textbook comparisons, as well, to give us sort of a descriptive analysis relative to what is algebra. That will be fueled by the research that some of the other groups are working on; particularly the Learning Processes group as they move into algebra itself.

We have, and it's matter of public record, created a listing of major topics of school algebra that will be supported with not only the major topics, but a discussion of those topics in prose, supported by research, hopefully to be made available at our next meeting in Miami. And then there will be an appendix that will take that relatively brief discussion of algebra and expand that to a full elaboration of algebra.

The corollary to the question relative to the definition of algebra is the question, “What are the essential, foundational concepts and skills that lead to algebra?" Again, there's an analysis here. Our analysis is looking at the mathematics taught in grades K-8 in top performing Trends in International Math and Science Study (TIMSS) countries.

We are also looking at the differences in curriculum approaches in those top-performing countries. We have looked at the NCTM Curriculum Focal Points. We are also looking at the mathematics skills and concepts in the six highest rated state curriculum frameworks, and also a yet to be completed survey of teachers of algebra in this country. The survey is going to begin very soon.

So we will come out of that with a draft of the foundations, the essentials that students ought to have prior to experiences in algebra. This would not be an entire full curriculum, but the elements, the critical foundational pieces that lead to algebra. There will be a discussion of those as well and an elaboration. You can perhaps see the analogy to the algebra piece here as well.

A third question, does the sequence of mathematics topics at grade levels prior to algebra affect algebra achievement? For this final question we have a work in progress in this area. We're intending to look at the following: programmatic research on recently developed curricula, benefits of an integrated approach and the role of integrated mathematics in this whole configuration of school mathematics, particularly algebra at the secondary level, and the research on the placement of algebra. By that I mean the actual grade placement of algebra, the percentages of eighth grade kids taking formal algebra, or for that matter, lower than grade eight. So that's an analysis that we've begun as well. And that's where we are.

VICE CHAIRPERSON BENBOW: Are there any questions? Well, hearing none, Skip, your coffee is up here. All right, if I could now have a report from the Learning Processes Task Group. Dave Geary is the Chair of this task group and he's going to be delivering the report.

DR. GEARY: Do I turn on the timer? No, all right. This will be short anyway. Contributing members to this group are myself, Dan Berch, Wade Boykin, Susan Embretson, Valerie Reyna, Bob Siegler, and Jennifer Graban is the Department of Education staff member assisting us.

As you know, last time we presented a detailed review of what we had done at that point covering basic principles of learning in cognition, mathematical knowledge children bring to school and math learning in whole number arithmetic. So I won't bore you again with those details other than to remind you that is what has been completed.

The other groups reviewed that work this time and we found the comments to be very helpful and suggestions for our revisions to be very helpful. Between now and our next meeting in six weeks, we're going to take these comments and suggestions into consideration and revise these three sections accordingly and hopefully bring it up to something very close to a final draft. As part of those revisions we will begin to extract out policy recommendations more explicitly in there, as part of the text and probably a separate summary section.

Between now and June we will also be working on a drafted section of the social motivational affective processes. We hope to have a nearly complete section of that to be included with the other three sections, and the revisions for your review at that time. We hope to have those sections completed after the June meeting, nearly finalized.

Between our June meeting and the meeting in St. Louis in September we will complete the sections on fractions, estimation, geometry and algebra. The latter two areas may have less work than the other areas, but nonetheless, we will review that and point out areas where there are substantial holes.

For the St. Louis meeting, as well, we hope to review differences and similarities across race, ethnicity, socio-economic status, and gender in the key areas that are included in this report. We will also have a section on recent work in the brain sciences in math learning and mathematics cognition. Of course, we will also take comments and recommendations at that meeting and we will hope to have all of those changes done by the October meeting, to have our section of the report complete by then. And of course, during all of these revisions we will be working on integrating our aspects of the report with the aspects of the other four sub-groups. That's it.

VICE CHAIRPERSON BENBOW: Are there any questions? Seeing none, hearing none, thank you. All right, at this point in time we will move up with a presentation from the Instructional Practices Task Group. Russ Gersten chairs the Instructional Practices Task Group, and Tom Loveless and Joan Ferrini-Mundy will be joining him to present our work so far.

DR. GERSTEN: Myself, Camilla Benbow, Doug Clements, Bert Fristedt, Tom Loveless, Vern Williams, Joan Ferrini-Mundy and Diane Jones are members of the group and Marian Banfield is our Department of Education support person, team member.

Quickly, I just want to review again, we shared this at the last meeting, but we have basically firmed this up a little bit. The core of our report on each of the six topics we've agreed to look at, with the possibility of a seventh if time permits, will be experimentally high quality, quasi-experimental studies using criteria very similar to the What Works Clearing House. I'm not going to go into the technical details now. We've had an excellent team from Abt Associates and have worked collaboratively and productively with them.

Other studies that we will look at and use to inform our interpretation of the findings, our framing of the issues, and our thoughts about future research include any other type of quantitative studies, descriptive or correllational studies, qualitative and K studies. We also have a group of tier-four studies that are flawed experiments, studies that have some level of serious problems with them. We will only mention them with extreme caveats, because these are the ones that the data is really not interpretable due to the serious types of problems. And again, the details of this are flushed out in our preliminary writings.

Tom, Joan and I are going to share just where we are in the first three topics and we'll start with Tom.

DR. LOVELESS: Thank you, Russell. As Russell pointed out, Abt Associates performed a meta-analysis for us. First they conducted a search of the literature applying the criteria that Russell described. The first topic that we wanted to look at was the whole issue of student-centered learning versus teacher-centered learning, considering that as a continuum.

Within that literature the search produced over 100 studies. I can't remember the exact number. I think it was 126. And what we did was then apply our criteria, which screened down the literature. Of the remaining studies we then grouped them by their common approach or intervention that was tested. The one area that leapt out as having a sufficient number of studies to really draw some conclusions about was cooperative learning and peer assisted learning. And those are the results I'd like to show you.

First of all, in cooperative learning one technique that was studied was team-assisted individualization. This is an intervention that involves grouping students into groups of four or five and then giving the students work on particular areas in which they have shown deficiencies. And then the students work as a team for a period of time, as opposed to say doing individual seatwork. And then the students are tested, both pre and post tested.

In these particular studies, these are all tier-one studies that we're looking at. The students were randomly assigned to both treatment and control groups. As you can see in terms of math concepts the effect was trivial.

In math computation, however, there were six studies that produced seven pooled effect sizes. The pooled effect size is .340, which is statistically significant. You can see the p-value, .002. So this particular finding is actually the most robust finding that we came up with.

I want to caution right up front that this does not mean that simply putting students into groups and then giving them math to do, necessarily produces results. These are highly structured interventions. They are not simply testing grouping, but they're testing a particular form of grouping with a specified award structure.

The second area in which we found sufficient research to perform a meta-analysis was student teams achievement division. This is another Johns Hopkins invented intervention. And we found no significant effect.

In terms of peer assisted learning, again, we found an effect on computation. This is one particular study, just one study. It had classroom level data, where classrooms were randomly assigned to treatment and control. Lynn Fuchs was head of the research team. We found a significant effect there in this particular study of 0.441. Most researchers would consider that to be a modest effect. And the p-value of .021 shows that it is statistically significant.

We called this next group of studies “other cooperative learning strategies” because they didn't fall under any particular definition of the cooperative learning strategy, but they did test cooperative learning. The Mevarech study, for example, is out of Israel, and the effects size of .230 is also statistically significant. In this particular study the students were assigned in pairs to a computer-assisted learning intervention.

So in one intervention, students worked individually at the computer and received their math instruction. In the experimental condition the students worked in pairs at the computer and received their instruction that way.

Finally, we call these mixed approached

and interpret them with some caution because not only was either peer assisted learning or cooperative learning part of the intervention, but there were other characteristics of the intervention. Other things were modified. Curriculum was changed or something else was going on as well as peer assisted learning.

So we can't isolate cooperative learning or peer assisted learning and say that was the thing that produces this positive effect, but they should be noted. Busato was a study out of the Netherlands. And that's a large effect, the largest of the studies that we looked at here, .634, and that is statistically significant. This is another Fuchs study of peer assisted learning and I talked about that earlier.

DR. GERSTEN: I'm going to talk a little about the work on formative assessment. We actually found a set of high quality studies. The first question is, does it help students? Is math achievement raised if teachers weekly, every other week, have some assessment of where kids are, what they've learned or not learned and some valid measure.

And the second one is for teachers to get the raw data and to try to make sense of it and develop instructional plans. And the other things, which we call enhancements, are giving specific tools or strategies or procedures to teachers to help them figure out how to use the data, what they might do with it. So those are our two research questions.

We found ten high quality studies, which is a lot for most topics in education. This would not necessarily, in medicine or public health, be a lot, but for education ten is a lot of this quality. All are in the elementary grades. The measures are both concepts measures and computation measures, very similar to what Tom showed you.

The technical characteristics of the measures seemed fine. But the content validity, we're having three experts on our Panel review that in a bit. And that is not completed yet.

This is the type of formative assessment that was done in these particular studies. It isn't the only way to do it, but it is the way it was done in this set of ten studies. Basically a sample for the year's state standards, the kinds of things kids are supposed to know by the end of the year by May or June, were used to generate items. And each of these tests given usually every other week, typically on the computer, kind of take random samples of the items.

So this is very, very different than the way formative assessment is done in most classrooms in the U.S. or around the world. The idea is that this way you can really track growth. They wind up just in terms of psychometric and technical qualities to be far superior to the typical weekly unit tests. Because you also get at not just what the kid learned during the week, but what they retained, and their ability to use what they already know to figure out stuff that might come out in the later half of the year. So it winds up working better. There may be other approaches, but we just don't have the level of evidence on other approaches.

There is a consistent statistically significant effect for teachers (using basically random assignment, high quality designs). Use of formative assessment does raise student achievement by approximately a quarter of a standard deviation or ten percentile points, which is not too bad on the fact that it's repeated or replicated again and again.

The second thing in terms of these enhancements is the effect more or less is doubled, and I'll show you in a second what the enhancements are. When you look at the whole set of them, what you need to do to do statistical tests, you get a sense that the effect is about double, so it gets closer to 18, 19 percentile points.

The only thing is, these studies of enhancement were almost all, with one exception, done with special education students. So that is something to keep in mind.

In one study after the performance data was analyzed, these enhancements are basically the computer-generated practice, which became the basis of tutoring sessions. So kids were getting help on material they needed help in.

In another study, the teachers didn't get their hands on materials but had a sense for each child and for the whole class. These are areas that the kids need help in. So again, it was a way to guide time for differentiated or individualized instruction.

In one case there was a bank of experts, math coaches, math specialists who developed ideas when kids are having trouble with place value and hundredths and thousandths. Again, this is a way to intensely work with a small group of kids.

And the last one was kids learn to monitor their own progress. They themselves can see how they're doing and figure out what are the areas they need help in.

So that is where this stands. But it is actually a pretty solid basis for making recommendations in our view and we've got an input, which we will incorporate, from the other groups.

DR. FERINNI-MUNDY: The third category where we've made some progress is in the area of “real world problem solving.” And the reason that we've put that in quotes -- we have found is that this notion of "real world" problems is not an unfamiliar idea -- curriculum and it's been available --

VICE CHAIRPERSON BENBOW: Could you speak into the mic --

DR. FERINNI-MUNDY: Sorry, sorry. Many current policy documents call for the use of real world problems in mathematics instruction and this is reflected in some instructional materials as well.

Now the reason that "real world" is in quotation marks is summarized here, and there is a discussion of this in our draft material thus far coming from the literature.

One of the issues with this topic is that real world is an under-specified construct. We have a variety of meanings that appear in the research, that appear in the discussion by developers about what they intend with this. And we've listed here just a few of the areas that we're seeing come up in the descriptions of what people mean by this general area.

So you see, for example, literature that discusses real world problems as problems that would be meaningful, appealing and motivating for students from contexts that they know, from imaginary situations, from mathematics. Sometimes the discussion focuses more on what are called authentic problems. That would be similar to those in applications beyond the school setting. Often there is description of such problems as being complex with multiple steps and involving integration of concepts. The idea of open-ended problems, problems both with multiple solutions and possibly multiple solution paths are included sometimes in these descriptions.

We also are finding in the literature that there are many arguments from a variety of places based on beliefs, experience and research, both for and against the various types of real world problem emphases that you've seen in the previous slide.

This makes it complicated to review the research, and at this moment we're looking at only 12 studies that Abt has located for us through their searching. Three of these are quasi-experimental studies that have examined the impact of what I would call full-blown curricula that feature, in some sense, a real world emphasis. And these studies all have methodological issues, but they are providing us with some insights and some ways of framing this discussion that will be very helpful.

There are nine other studies that we have found that look at the impact of various types of instruction using "real world problems" and/or instructional strategies that are meant to help students solve real world problems. And again, these studies have methodological issues but they're raising important conceptual issues for our discussion. So we are in process with this but we wanted to let you know where we are with it at this stage.

VICE CHAIRPERSON BENBOW: Do we have any questions? Russ?

DR. WHITEHURST: More in the form of suggestions/questions. Tom, as I looked at the presentation there seemed to be occasions where you would pool effect sizes across a group of studies and other cases where you simply highlighted a positive effect size for one study and left uncommented upon smaller effect sizes for other studies. So, at some point that needs to be rationalized.

DR. LOVELESS: If I could just respond?

DR. WHITEHURST: Sure.

DR. LOVELESS: We pooled when it was clear the intervention was similar across the studies. In the ones that we did not pool, we did not pool them because the opposite was true. It looked as if there were key parts of the interventions that just differentiated them.

VICE CHAIRPERSON BENBOW: I think we need to stop. Can you hold those questions? We can pick that up. I think we hear the bells. I say since we stopped and we'll pick up this dialog and discussion back and forth, but since we've stopped and we're so close to 11:00 o'clock, we have one more minute, let's just wait, we'll pick it up.

DR. LOVELESS: Can we leave now?

VICE CHAIRPERSON BENBOW: I suspect there might be more questions coming, Tom.

DR. LOVELESS: I'm practicing the button pushing.

(Whereupon, a short break was taken for a message from the principal regarding the Illinois day of mourning for Virginia Tech.)

VICE CHAIRPERSON BENBOW: All right, Russ, if you want to pick up where we left off?

DR. WHITEHURST: My other question or point or suggestion is that as a Panel, I think we need to be cautious or perhaps come to some understanding, shared understandings as we're talking about small, medium and large effects. There is nothing out there that anchors those terms, an effect that might be considered small, it could be large if it accumulated. Something that's an effect over a two-year period would have a very different meaning than an effect over a two-week period. Thanks.

DR. GERSTEN: That is one issue that we're grappling with and we're going to be working with Mark Lipsey on as he has some time for our group. And it's an excellent point and one that, guidance from any members of the Panel, Institute for Education Sciences, et cetera, would be really appreciated.

VICE CHAIRPERSON BENBOW: Tom, did you have a response that you wanted to make? I certainly didn't catch it.

DR. LOVELESS: No, I agree totally.

VICE CHAIRPERSON BENBOW: Wade?

DR. BOYKIN: Yes, with regard to these 12 studies on this slide here, you recognize they all have flaws methodologically, but are there any kinds of tentative inferences you can draw from these particular studies?

DR. FERINNI-MUNDY: Actually, we're still really working on that. It's a little bit early. We have to decide whether these flaws outweigh what we actually might be able to say.

Part of the issue has to do with the outcome measures, which vary greatly on these kinds of studies. And some of them will feature only items that are aimed at testing students' ability to solve “real world problems.” Others are more standardized measures that include a range of outcomes. So I think it's a mix of having our mathematician experts take a look at these outcome measures so that we can say something about what the results would mean. So we're really mid process on that one.

VICE CHAIRPERSON BENBOW: Bob?

DR. SIEGLER: I'd like to ask Russell a question about the formative assessment work that you talked about.

If I understood it right, kids are not only getting instruction, the teachers are getting information, but also the computer program in some or all of the studies is generating problems that are designed to remedy the children's learning difficulties. Was that a misunderstanding on my part?

DR. GERSTEN: Bob, that's only the case in several of the enhancement studies. So when we look at the whole set of ten there is a condition where the teachers and sometimes the kids get the numbers, but that's it. They get the feedback. The enhancement studies, that smaller set with the special education students, is where they get, in most cases some additional, either information for instruction or additional specific ideas for how to teach the kids.

Is that clearer?

DR. SIEGLER: Yes. You might want to consider the older literature on adaptive computer assisted instruction as another way of thinking about formative assessment, because here it isn't the teacher that is getting the formative information but rather the computer program is getting it for itself and generating problems on the basis of that.

DR. GERSTEN: Those studies didn't come up in the search. I think some ideas and leads on those, I'm dimly familiar with them, but I think they would be appreciated and we could look at those. We can talk to Abt about expanding the search to look at those.

VICE CHAIRPERSON BENBOW: Liping?

DR. MA: Yes, do you have any research available about the relationship between real world problem and regular world problem?

DR. FERINNI-MUNDY: We have research studies in both areas that we're looking at, but I don't recall that we have any that actually looked at the relationship between the two. So if you know of something or if others do, that would be helpful to us.

DR. MA: Thank you.

VICE CHAIRPERSON BENBOW: Valerie?

DR. REYNA: Thank you. I have a question about tier-three evidence and just what you're thinking was. I should say at the outset that that level of evidence, qualitative research is certainly a valid scientific method.

That having been said, the question for your group in particular is really about efficacy, I would think, instructional practices, that the question ultimately is one of efficacy. So, what was your thinking about inferences from samples to populations or to questions of efficacy from tier-three level research, as you characterize it?

DR. GERSTEN: That's something we've discussed and thought a lot about. We do not exhaustively review tier-three studies. But if there is a study, and it's based on either the Panel's judgment or the author's judgment, that helps us frame an issue or interpret findings or interpret findings that are erratic. So it's only used to aid but there are no results emanating from those studies. Definitely the ideas and concepts there can be used for ideas for future research or to help us frame current understandings of issues.

DR. REYNA: So you're saying you're using them for theoretical purposes? And would there be any sense of which evidence should necessarily bear on theory?

DR. GERSTEN: Evidence, well, so we're using them if they help understand a phenomena or a pattern or finding.

DR. REYNA: So you're saying that qualitative research allows you to infer causal mechanism?

DR. GERSTEN: If there are ideas in the published literature that help us understand phenomena, that’s helpful. So that's what they're used as, as basically sources for ideas.

DR. REYNA: I won't continue the debate, but what I'm saying is, therefore, this would be a source of speculative opinion and it would be marked as such?

DR. LOVELESS: Yes, it would be marked as such and it would be used to generate future hypotheses. For instance, in the cooperative learning field we have this effect, this effect that's statistically significant.

We may want to propose, and we haven't gotten to this point, but we may want to propose future hypotheses that could be tested as to what are the mechanisms of this intervention that are generating this positive effect. And the tier-three studies could help us frame those hypotheses. It should be clearly labeled as not somehow causally verified in the literature.

VICE CHAIRPERSON BENBOW: Do we have any more questions? Well, seeing and hearing none, thank you. We'll move on to the next group, teachers and teacher development.

DR. LOEWENBERG BALL: Okay, I'm reporting on behalf of the Teachers Task Group. The names of the members of this group are on the slide and Ken Thompson is the staffer with our group who's done a great deal of work to help us.

So first, I just wanted to review for all of you what the four questions are that the task group is considering. We will only be reporting on question one at this meeting.

The first question has to do with the relationship between teachers' mathematical knowledge and their students' achievement.

There are subsequent questions that we've begun to work on and that you'll hear about at upcoming meetings that include what is known about programs that help to increase teachers' knowledge, both pre-service and in-service. It also includes the relationship of what teachers learn in those programs, evidence about what they in fact learn, and the relationship to, in particular, their students' achievement as a result of their opportunities to learn.

One of the other areas has to do with elementary math specialists. What we've been able to determine so far is that we won't be uncovering studies that link, as the question asks, the effectiveness of math specialists programs or math specialist staffing to student achievement. We will go ahead and begin to explore what the range of models is that exists out there, what the differences are among them and what's known about what kinds of qualifications are used to place people into such roles. We'll also be looking internationally to understand the ways in which a math specialist may be employed in other countries.

And finally, we'll be looking at what's known about strategies for recruiting and retaining really highly qualified, skilled teachers in teaching mathematics. Both of these last two areas will probably turn out a little bit differently than our first two. For example, in question four we'll have to look at data and research beyond specifically mathematics teaching, to understand what's known, in particular, about the recruitment and retention of teachers in general.

Question one is the one that looks at the relationship between teachers' mathematical knowledge and their students' achievement. So our group thought it would be useful just to reiterate for ourselves why this is such as important question for the Panel. And we saw three essential reasons.

There is substantial research and anecdotal evidence that U.S. teachers’ levels of mathematical knowledge are often too low for the work they're being asked to do. That is, they don't know math deeply or well enough.

There are many ways people describe this. There is both robust research evidence on this and plenty of anecdotes floating around. And our charge was not to try to trace the documentation of that weakness, but it is what compels this question.

We also wanted to note that there's an increasing trend of increased requirements for American students to take more mathematics. So for example, in my own state, Michigan, where we've just moved to a requirement whereby all students will take four years of high school mathematics and that high school mathematics is actually shaped at the state level. You can see the increasing need to have qualified teachers who can deliver that content to a wider range of students than ever before.

And finally, we'll say a little bit more about this. There are some critical areas that we're going to try to display still more than we have already, in which there is a significant need for qualified teachers to be teaching. And let me just show you briefly what those are.

One is to look at the likelihood that a minority student or a student living in poverty will have a teacher who's either certified in mathematics or has a major or minor in the field. Look at this chart taken from the 2003 Condition of Education Report. They are not as recent data as you might like, but I think it helps to exemplify the problem. You can see that minority students or students living in poverty have roughly twice as high probability of having a teacher who does not hold a major or minor in the field or isn't certified in mathematics. You can see that only science has a situation that's about that dramatic.

Another way to think about it is to look at the particular problem of high school and middle school teaching. This graph shows percentages of middle school and high school students who have mathematics teachers who are qualified by either of these criteria, hold a major or minor in the field or are certified in mathematics.

You can see mathematics really sticking

out up there, that dark red bar. This represents middle schools students. So roughly one in four middle school students is being taught by a teacher right now who does not hold either of those ways of being qualified to teach mathematics, and even one in ten at the high school level. So these seem to us to be critical reasons to highlight this in our report.

So what does that mean we might need to know to inform policy better? One of the basic questions that question was about is, how does teachers' mathematical knowledge relate to students' learning. But more than that to inform good policy we would have to know how much mathematics do teachers need to know to be effective and what mathematics do they have to know, and in what ways. And you can see why those subsequent two questions matter, because simply knowing the teachers' knowledge in mathematics relates to student achievement, which is something everybody already believes anyway, doesn't provide all that helpful guidance for cost effective and effective interventions to improve and increase teachers' mathematical knowledge.

If you think about this question about what is the relationship is between teachers' mathematical knowledge and their students' achievement, there are two basic methodological issues.

One is how would you measure teachers' mathematical knowledge and what would you mean by students' mathematical achievements? I just want to briefly say how these two things are treated so far.

So for measures of math teachers' mathematical knowledge, in our review of the literature we're looking at three different ways of measuring teachers' knowledge.

One is teacher certification in mathematics that is indirectly also the result of a test. But it's separate in the studies from the second type, which looks at teachers' educational attainment in mathematics measured either by their degree, a degree in mathematics or levels of course taking.

And then we have what we are calling currently more direct measures, that is measures of teachers' mathematics of the curriculum they have to teach or of the content of their level or beyond. There's not as much research in this area, but we consider these to be less indirect than the first two.

And then the question is, what about students' mathematics achievement? How might researchers examine this? The studies that we selected and considered to be high enough quality were longitudinal data on students' performance using pre-test controls. So that what we're looking at essentially are either longitudinal growth models, gained scores or some form of covariant adjustment models. These are not cross sectional studies that we're using.

Further, actually one would want to know about what the outcomes are in gained scores and about how teachers' mathematics impacts the instruction that students receive and thereby the learning that they are able to accomplish. Our studies don't have that sort of measure and ideally we need more research that traces this a bit more closely inside the so-called black box of instruction so that we could make better policy decisions.

Now I'm just going to report briefly what we've learned in those three ways of measuring teacher knowledge, teacher certification, course work and direct measures. Looking at the effect of teacher certification in mathematics on student achievement, there are really three issues that we uncovered in the studies that we examined.

First of all, it's worth noticing that teacher certification is a pretty inexact measure of what teachers actually know. There are some substantial problems of selection bias in these studies. That is, it doesn't isolate the variables very well. There are other things that might coincide with certification that would make it difficult to say that what you're measuring alone or testing alone is certification.

And further, and we'll have to do a better job than we've been able to do so far to sort out the different things that actually are called certification. There isn't some uniform single thing that's being examined here.

But that said, in these studies the effect of teacher certification remains somewhat ambiguous. And maybe because of what I just said you can see why that might be methodologically.

Of the other studies that met our quality standards, four showed a positive effect of teacher certification on students' learning and four others showed no effect. Actually, it's not true that no results are significant. Some of the results were statistically significant.

There are some complications in these studies partly because of the inexactness of this as a measure of teachers' knowledge. But we consider this to be, nonetheless, a very important policy question since it is one of the ways that policymakers can intervene to assure qualified teachers. So we're going to be surveying what the specific certification requirements are, particularly at the middle school level since that's, given what I said earlier, a particular area we think we could say something about, and try to look a bit more closely at studies that may compare teachers with different kinds of certification.

So now I'm going to move onto what we were learning about teachers' mathematical study, which is the second of the ways that math knowledge could be measured.

Here we're looking at teachers' college level mathematics study. One issue is that course taking isn't a direct measure of what someone knows. For example, there would be a lag effect of what they learned a number of years ago and what they now know. It is not necessarily going to tell you what they know at the moment.

And furthermore, a different sort of problem is that these courses, particularly mathematics, may not correspond very closely to what it is that teachers actually teach. If you look at the content, for example, of the math major, many of the courses in the math major don't align all that closely with the content of the high school curriculum. So there may be some issues of selection bias here.

But that said, there's slightly stronger results here that are worth paying attention to. We see more consistent findings here than we do in those certification studies that I reported. Of the nine studies that met our criteria for high quality research, seven of those do show a positive impact of teachers' course taking or level of attainment on student achievement. One showed no impact and one showed negative impact.

It’s worth pointing out here, given the Panel's purview, that most of these focus on secondary school students and we did not uncover evidence that related teachers' course taking at the college level positively affecting student achievement at the elementary level. And further, again to say, we don't know very much about what these courses are about. So these are still somewhat distal from teachers' content knowledge. But still, we're seeing that there's a positive effect, which is a signal in the direction that many people would believe anyway.

So the third area has to do with studies in which there is more direct measure of teacher knowledge using tests of some form. While there may be closer estimates of what teachers actually know, overcoming some of the problems I pointed to in the other two areas is still an issue.

Some of these measures haven't been validated so it's not obvious what we should make of measures invented by researchers to study teacher knowledge. We also found extraordinarily few studies of this kind.

Still, the numbers of studies that met our criteria allowed us to say some things about what we were learning. We had eight studies, five of which met our standards. Two of those showed positive effects that were significant. One showed positive effects, although not statistically significant. And two found more ambiguous results.

So we think that, generally we feel supported in saying that hereto, there is support for the notion that teachers' mathematical knowledge having a positive impact on students' achievement.

So, if we were to make two tentative claims at this point I think there are the two we feel that we can say so far, that "knowing" mathematics is likely a significant factor in teaching effectively. Now, you might say, well, you didn't have to do all of this to come up with that.

I'll say a little bit more about what we think will be needed in order to be able to do something policy-wise with that. And "knowing" is in quotes here precisely because what I've been saying over and over is, none of these gets very close to the notion of what exactly does somebody have to know and in what way to teach well. We think that given the scale problem of how many teachers we're talking about it would be useful, from a policy perspective, to know exactly where to target, how to increase and improve teachers' knowledge.

It’s worth underscoring again that the notion of college level study may predict effectiveness for secondary school teaching, but we did not define that for elementary. So given the critical nature of the need to improve teachers' content knowledge at the elementary and middle school level, this gap in the research does suggest that we have a need to probe more closely into the mathematical knowledge needs for the K-8 level.

We don't know, as I've been saying, enough about what teachers actually have to know. We don't know quite enough about how teachers' knowledge affects the quality of students' learning. That is, how it interacts with instructional practices, for example, or with teachers' knowledge of learning to enable them to actually effectively address what students produce in class or to design instruction.

We also don't know enough how much course work makes a difference at different levels of schooling. If you think of the practical implications either through assessments or course requirements, one will need to know that simply saying, well, let's just have everybody take a major, clearly isn't supported by what we've been able to see so far.

So we wanted to end by saying what we think might be needed to inform policy better. We think there should be investments in better and more reliable, more proximal measures of teachers' actual mathematical knowledge. We need a better way to understand how the teaching of mathematics demands mathematical knowledge so that we can target the research we're doing in a more focused way on the actual mathematical demands of the work. We'd like to see studies that had better designs that would permit stronger causal inferences that, for example, overcame some of the problems we uncovered. We need studies that do a better job of isolating variables, overcoming selection bias, and looking more closely at the impacts on instruction.

So in terms of just telling you where we're going next, we'll be trying to gather more detail about certification requirements including not only what's required to get certification, but what the assessments, what the cut scores are and the nature of mathematics asked on some of those tests. We're going to look more closely, as I said, at teacher qualifications at the middle school level. We’ll be trying to compare effects of different forms of certification and we'll try to improve the way we've consolidated our estimates of effects across the studies that we've examined.

That's really the detail of our report on question one. You'll be hearing more about questions three and four at the June meeting and question two, I think, by the following meeting. Question two, again, is the programmatic intervention question.

VICE CHAIRPERSON BENBOW: Since I'm controlling the microphone, let me go with the first question. Could you clarify for me, when you say the evidence regarding elementary math, is it that there is no evidence assessing the importance of mathematics for our elementary teachers' effectiveness? Or is it that there is evidence that is showing no effect?

DR. LOEWENBERG BALL: No, no, I didn't mean to say anything that broad. It's that the course work studies don't show an effect of course work on teachers. But in the third grouping of direct measures, one of the high quality studies showed a significant effect at the first grade and third grade level of teacher's mathematical knowledge. If I said we don't know that it makes a difference at elementary, I didn't mean to say that. I meant that the course work studies don't show us that.

VICE CHAIRPERSON BENBOW: Thank you. Doug?

DR. CLEMENTS: Your third question about math specialists speaks to more instructional effects rather than pure math knowledge. So I was wondering, as a newcomer to the Panel, if you guys had discussions similar to Shulman's seminal pedagogical content knowledge versus content knowledge and whether you were going to even try to look at the former?

DR. LOEWENBERG BALL: Well, we haven't yet found studies that examine the effects of math specialists on anything. So if we were to design studies we'd presumably want to know how having a math specialist effects instructional practice or student learning. But we're not finding that.

DR. CLEMENTS: I'm sorry, all I meant is that it looked like you were looking at instructional kind of issues, but most of your presentation was about mathematics content knowledge, right? Not about pedagogical content knowledge, which I know you've contributed to that literature. So I was interested as to whether you thought that was just too difficult a problem to also address or there just being no studies or was it just the decision, if not looking at specific kinds of knowledge that are relevant instructionally?

DR. LOEWENBERG BALL: Yes, I think maybe what you're asking is, there are many kinds of knowledge that potentially influence teachers' effectiveness and this group appears to be focusing primarily on content knowledge, and that's true. We have been doing that. Although, as we look at the literature we're open to looking rather broadly at what is defined as mathematical knowledge. As it happens, pretty much all the studies with maybe one or two exceptions are looking at content knowledge measured rather narrowly.

It will have to be one of the recommendations of the Panel, however, to look more broadly even with our intersection with the Learning Processes group, which may help us to understand how knowledge of students' learning of math might affect teachers' effectiveness.

VICE CHAIRPERSON BENBOW: Vern and then Wilfried?

DR. WILLIAMS: You may have looked at this and possibly I missed it, but have you compared the number of math courses and the types of math courses taken by teachers in some of the high scoring countries, K-12, with teachers in our nation?

DR. LOEWENBERG BALL: Looking at, in other words, requirements, and qualifications to teach at different levels. We should do that. We haven't done that yet. We've just begun now to look at certification requirements. Initially we were looking at studies that looked at effects. I think we should broaden to look internationally. We've known we should do that with math specialists and this would be a close cousin, I think.

VICE CHAIRPERSON BENBOW: Wilfried?

DR. SCHMID: Two questions. First of all, have you or will you look at, let's say, the effects of professional development? Or will you do this only indirectly through, let's say, assessing content knowledge of teachers, which might or might not be imparted by professional development?

DR. LOEWENBERG BALL: Our second question, which we have not yet reported on will be examining programs at the pre-service level and professional development intended to increase teachers' mathematical knowledge. And we'll be looking to see whether and how they affect increases in teachers' knowledge and their effectiveness. Is that what you mean? Yes, we'll be looking at that directly. We just haven't yet.

DR. SCHMID: In the second question, the question of retention of teachers, will you look at the question of differential pay for mathematics teachers?

DR. LOEWENBERG BALL: Yes, absolutely. So teacher pay will be one of the aspects. We'll have a grouping of different possible strategies for that and teacher pay will be one of those.

DR. SCHMID: Well, I said, differential pay meaning --

DR. LOEWENBERG BALL: A different pay for math teachers.

DR. SCHMID: -- different incentives specifically targeted to mathematics teachers.

DR. LOEWENBERG BALL: Yes, you said it right and I said it wrong. That's what we would be looking for.

VICE CHAIRPERSON BENBOW: Tom?

DR. LOVELESS: Have you done anything with these studies to pool their effects, applied any analytic techniques so that we could get an idea of the size of the effects overall and whether or not they're statistically significant?

DR. LOEWENBERG BALL: This has been one of the challenges that we're still engaged in trying to find a way to do that. So, I saw the Abt staff shaking their heads back there, not to say no, but reminding us of this challenge. We have been looking for a way to do that, that we think makes sense.

DR. LOVELESS: And then secondly, on the literature addressing the question about college level courses, do those studies differentiate between courses taken in math education as opposed to mathematics departments. It was mentioned earlier about international patterns on this and the United States really is an outlier in terms of our eighth grade algebra teachers. At least TIMMS show most of our algebra teachers in eighth grade received their math education in a school of education. And around the world, most eighth grade algebra teachers received their education in math departments. So I'm wondering if the studies allow you to take a look at where those math courses were housed?

DR. LOEWENBERG BALL: I think there's several questions embedded in yours. One is where the math courses are housed. Another one, which you may not have meant to ask, but needs to get asked, is that some of these studies also look at the effects of course taking and math methods. And interestingly, those sometimes show stronger effects on student achievement than the pure math content. So we haven't broken that out well.

And third, we are going to look very closely at middle school requirements. So it's sort of circling around your question. These studies don't necessarily tell us where they're taking them. But your question is pointing to something that we'll try to get into in several different ways.

VICE CHAIRPERSON BENBOW: Sandra?

DR. STOTSKY: I know that you haven't reached the second question yet. Do you have any sense now whether any of these studies will be looking at the pre-service programs, student teacher issues, placement issues for student teaching, the evaluations that are done as part of student teaching and what those look at in relation to mathematics knowledge as opposed to mathematics teaching? This is before a prospective teacher exits --

DR. LOEWENBERG BALL: So part of what I think your question might point at is how deeply we're going to go into learning about what programs actually are, as opposed to simply looking at programs and whether they have affects. And I think that isn't a question the group has had a chance to discuss yet, but I think it would be helpful, in the same way that looking at certification requirements for some of these details helps with question one. Because I suspect, as you do, that there will be clues there about the preparation that would be useful to uncover. We'll have to see. I think we'll just have to see what we can do with that.

VICE CHAIRPERSON BENBOW: Are there any more questions? If not, thank you. Oh, sorry.

DR. WU: What I believe part of the next steps for question number one is to pin down the nature of the knowledge teachers need to teach. I think that's one of the main issues.

DR. LOEWENBERG BALL: Yes, so Wu's talking about something we began to discuss here, is how far our group may go into actually making some hypotheses based on our judgment and what we've read about the correct answer to the what question. Since the studies don't uncover all that much about the what, how far might we go at least framing what we think to be reasonable hypotheses about that. And you're right, that's one of our issues. That's why you get to ask a question.

VICE CHAIRMAN BENBOW: Any more questions? Okay, thank you. Now we'll have a report from the Assessment Task Group, and giving that report will be Susan Embretson.

DR. EMBRETSON: Well, it was mentioned earlier that this was the first meeting of the Assessment Task Group. This is a slide showing the members who were present. Actually, we have another member who was present today and will probably meet with us, Douglas Clements.

Our main charge, since this was our first meeting, is to determine what kind of research question we want to look at, or research questions, plural. And we determined that really most important was a single question with many different aspects. And that question is, to determine the correspondence of National Assessment of Educational Progress (NAEP) fourth and eighth grade tests to selected state accountability tests for validity in assessing mathematics proficiency.

Now we limit the comparison to fourth and eighth grades because that is where NAEP is available, and NAEP, of course, is regarded as the national test.

Now when we look at validity, four aspects are particularly relevant for comparing NAEP to the state accountability tests. These four aspects are content validity, substantive validity, consequential validity and generalizability.

Let me say a little bit about what those kinds of validities are for those of you who are not familiar with that. Content validity is probably what you're most familiar with in educational tests. It's representing the content of mathematics. Tests are constructed with blueprints, which outline topic areas and their relative representation. That goes toward the content validity of the test.

Now, the content validity of the test needs to be made clear so people can compare it to say, some idea they have about what should be in that content.

The substantive aspect has to do with the underlying processes and theory about what is going into solving the test items. This kind of area might make contact more with our Learning Processes sub-panel or also helps define the nature of what is tested by particular items. Items can be formed on the same content topic in many different ways and, hence, can involve different processes required by the students.

Consequential validity is the impact of the test on defined groups of people such as gender, racial ethnicity, English as a second language, or disabilities.

Finally, generalizability looks at the impact of some features of testing that may impact score levels. Such as whether or not the test was presented by computer or paper and pencil, whether or not the questions are given in multiple-choice format, a constructive format, so on like that.

So these are the aspects that we think are relevant to look at when comparing NAEP to the state accountability tests.

Now what are the possible differences between NAEP and the state tests? Well, first the content may be weighted very differently between and within strands. And we don't know at this point without looking at the blueprints for NAEP and for the state tests just how the different content is represented. Secondly, the cognitive complexity of items may vary between NAEP and state tests. The same content can be tested by items that vary substantially in complexity. If you put in an extra sub-goal or extra vocabulary, for instance, it becomes a more complex thing. So there are many things that go into cognitive complexity. And I think this is one of the reasons why items, which measure the same content, differ in difficulty.

Now three, we want to look at the empirical difficulty of items that measure the same content. The distribution of the difficulties, again, may vary substantially between NAEP and the state accountability tests.

A fourth thing to look at is tool inclusion, calculators in particular, but also perhaps manipulative materials. These, again, may have impact on assessed proficiency levels and affect test validity, certainly generalizability, but also the substantive aspect of validity. When you think about calculator use, it may change the processes people are employing to solve the problem.

A fifth thing to look at is test delivery mode. We particularly want to look at differences between computer based versus paper and pencil tests, keeping in mind that computer-based tests are going to increase in popularity as time goes on.

A sixth thing to look at is the representation of items on NAEP versus the state tests. And of course, we have item formats ranging from true/false, multiple choice, constructed response, word problems and so forth. Now what we want to look at, of course, is how that impacts proficiency as well.

So, our actual comparison variables are pretty much what I just outlined. We're going to look at proportional representations of content from test blueprints, and cognitive complexity and conceptual skill level of actual items. This is probably going to be the most difficult because we have to have access to items to look at items and further, we have to have someone to do the looking. So this is going to be a most substantial effort. We certainly are not going to do this with every state. We are going to do a selected subset of states to look at this issue.

We also have not discussed what will go into the measures of cognitive complexity in the Panel and hopefully we'll discuss some of that today.

You would think the empirical item difficulty is easy to get, but it isn't. You're probably going to have to go on site and link particular items to difficulty. So that also needs to be done.

With regards to tool inclusion, there are some tests that allow tools and some that do not. So we want to look at that and try to understand what the impact is.

We will look at test delivery mode, which I've mentioned before, and item format, particularly as crossed with content. It could be that in some tests certain content is only measured by a particular format, and we want to understand better what impact that would have.

So, this gives you kind of an overall view of what we're looking at and how that's related to validity. The areas are proportional representation, complexity, and skill level, item difficulty, tool inclusion, test delivery mode, and item formats. Here are the areas that they are relevant to in validity.

I put down complexity and skill level under content because state blueprints do include cognitive complexity as part of their stratification. As far as I know there's not a great deal of satisfaction in many locations with this variable, but we want to look at that and see how that is determined and what their categories are.

That is also relevant to the substantive aspect of what is really behind the responses of students to items as well. I have also included item difficulty in a variety of locations; in particular, the generalizability is what comes most to mind. I think it also might be related to the substantive aspect of what's being measured.

Tool inclusion, calculators, as I mentioned, might change the nature of the problem solving process for the student and so we put that under generalizability and consequential validity.

And finally, item formats can also make a difference of the validity of a measurement. We should note consequential validity is everywhere. We would like to see the impact of varying representations of content, varying levels of skill, varying levels of item difficulty. I think this is most easily available from NAEP, but I don't know how easily available it is at the state level.

So this describes the groups that we're going to be looking at under the consequential validity aspect. So that's the conclusion. That's as far as we have gotten so far. Now we're going to determine how we're going to get this data.

VICE CHAIRPERSON BENBOW: Thank you. As you can see, we already have lots of questions. Wilfried?

DR. SCHMID: With the questions you have outlined, what kind of policy recommendations do you think you would potentially be able to make, depending on what you find?

DR. EMBRETSON: Well, that's hard to say, because we don't know what the status quo is on this relationship. I mean it could be that we'll look at the states and we'll think that the content there is better representative of some of the concerns of the other task groups in this Panel. Or we might not. We might like NAEP's content. I don't know what we're going to find. I'm also, for the first time, a member of the Learning Processes group and I've had some ideas of content that I want to look at, in particular, to see if it's represented on the different tests.

DR. SCHMID: The Department of Education has commissioned the study of the validity of NAEP specifically. I have been involved with that effort. It is clear to me that to study NAEP alone and to see how valid it is in the sense of phrasing questions correctly mathematically, having an alignment between the test questions and the NAEP framework, and having confidence that the methodology of the test questions is okay. These are very difficult questions to answer about NAEP alone. I am rather surprised if you propose to answer questions of this sort not just about NAEP, but also about several state tests.

DR. EMBRETSON: No, I don't think we're doing the same thing that you are. We're really looking more at the bigger picture especially in terms of content validity area and proportional representation of items. We're not going to check the reliability of categorization of items like you are in the various areas, nor to check their validity from the perspective of mathematical principles. So, no, we're going to do a more global analysis of this.

DR. SCHMID: But then again --

VICE CHAIRPERSON BENBOW: Last question, then I have to give some others a chance. Then we can come back to you.

DR. SCHMID: Well, I mean, it seems to me that from a point of view of actually coming up with policy recommendations we would need to know whether, first of all, NAEP is a reliable measure of whatever questions you ask about it. And then secondly, how well the state tests track that measure.

DR. EMBRETSON: Well, I agree with that. We can look at other features of NAEP as available in the literature. The standard areas around those scale scores and so on. We've already requested some data on that by level. However, you might mean reliability in a difference sense. I'm not sure if you have in mind the reliability of the categorizations of items but fulfilling the framework. That's a different sort or question.

VICE CHAIRPERSON BENBOW: Russ?

DR. WHITEHURST: Wilfried asked several good questions and raised several of the issues I wanted to raise. I still don't understand the purpose of the exercise you're engaged in. I don't understand how the analysis you intend to conduct will inform matters before the Panel. I'm not sure why the relationship between state NAEP, between state tests and NAEP is an issue. I'm not saying that it couldn't be an issue, but I don't think you've articulated what the issue is that you're addressing. At least I don't understand it.

To take a point that Wilfried has made and expand it a little, there are a number of studies going on in my office and in the department with respect to NAEP, to state NAEP correspondence. So it would be good, if you're not aware of them, to become familiar with them so that you don't try to do work that's duplicative.

It strikes me that another area perhaps that would be as interesting or more interesting would be the relationship between what's being assessed in this country, whether it's NAEP or state tests, which are most frequently based on the NAEP framework, versus international assessments. What do we know about what skill set it takes to be proficient on NAEP, how that would correspond to levels of proficiency that might occur in high-performing countries. And in fact, there are data that relate to NAEP standards to international standards that might be well worth exploring.

But to come back to the principle questions, it's just not really clear to me why you're doing what you're doing. I think that's probably a matter of explanation.

DR. EMBRETSON: Okay. I'm going to speak for myself. This panel has only met for three hours. So I certainly can't speak for that many people. But I believe that the state tests are closer to teaching than is NAEP because there are more consequences of the state tests.

And so by understanding, in fact, how proficiency is regarded at the state level, which is closer to teaching, I think we'll be in a better position to consider the factors that might be changed or recommended to be changed there and also on NAEP. NAEP is included in the mix because that's the national test.

VICE CHAIRPERSON BENBOW: Anybody from the committee like to add anything? I figured you would, Tom.

DR. LOVELESS: And this gets to Russ' question, the feeling of the task group was that when anyone hears the word assessment, certainly they think of NAEP. It's known as the national report card. If this National Math Panel issues a report that does not discuss NAEP it is in essence overlooking the most important test that the nation feels that it has that represents the United States' performance by students. So, it seems illogical for us to say anything about assessment without saying something about NAEP. Now we hope that the studies that are ongoing will provide some illumination of what these issues are.

The second reason, the second point I'd like to make is that we included state tests because those are the tests today in all 50 states, students take them in grades 3-8 as required by No Child Left Behind. Those, the results of those tests have consequences for schools. Schools are being held accountable for the results of those tests. So again, to issue a report from the National Math Panel without saying something about those assessments also seemed to us illogical. So that's the justification for NAEP and the state tests.

VICE CHAIRPERSON BENBOW: Okay, Russ, yes?

DR. WHITEHURST: Since Tom was looking at me during most of those comments. Let me just try to be clear, I think it is quite important for the National Math Panel to address what is being assessed with NAEP and what is being assessed at the state level. I think that's a critical issue.

What I was attempting to say is I didn't understand your framing of that problem and how it was going to generate answers that are relevant to policy concerns. To be specific, it would be of terrific interest, and I think you'll see an answer from my office before the Panel's final report, to find out what is the relative difficulty of state tests versus NAEP tests in defining proficiency. Or just have states set the bar higher or lower than is set on NAEP? And I just didn't see that kind of issue, which is kind of a natural policy issue emerging from the framework that had been provided.

DR. EMBRETSON: We understand that there's some data available on that as well that we can get from the Department of Education.

VICE CHAIRPERSON BENBOW: Let's move onto a couple more topics, Wu is next, Wade and then Bob.

DR. WU: Okay, I think what I want to ask really is in some sense colored by what Russ asked earlier. It seems to me that one of the main reasons for the founding of this Panel was that the student achievement in mathematics is behind international standards, levels. And so it seems to me that it would be good to look, we have to look at NAEP, yes, but the, one of the key questions is, if not the key question, is whether NAEP is measuring the right thing. Whether it's up to level.

And so it seems like, Russ says that you're doing the international comparison. And is our Panel going to duplicate or is it going to be a cooperative effort?

DR. LOVELESS: The department has conducted one study comparing NAEP and TIMMS, TIMMS of course being the international test. But I think Russ' point was a little different if I understood it right. They have national assessments and it would be comparing NAEP to those national assessments of other countries. And I'm unaware of any study that's done that.

DR. WU: I thought that's what Russ was saying, that you are doing that?

DR. LOVELESS: Russ was suggesting that we do that. We haven't discussed that yet. Again, we've only met three hours, so we're not really ready to issue a report yet.

DR. WU: Well, I don't ask for report. I ask for declaration of intention. But Russ says, can you confront him, I mean, maybe --

DR. LOVELESS: No, Russ' suggestion is very good and we're meeting this afternoon actually and --

VICE CHAIRPERSON BENBOW: You'll take it under advisement.

DR. WU: No, no, no, the issue is I want Russ to clarify. I thought you said your office was conducting an international comparison of NAEP with other countries' assessment. Is that right?

DR. WHITEHURST: What I said was that there are studies that, for example, draw conclusions about how many students in Singapore would be judged to be proficient on NAEP by cross walking results on international assessments. And I think those findings are pertinent and relevant to what the panel is considering.

VICE CHAIRPERSON BENBOW: All right, Wade?

DR. BOYKIN: I know it's early in the game for you all on your sub-panel, but in the executive order it explicitly says that this particular sub-panel should address the role of assessment in promoting math proficiency. I didn't quite hear that issue being addressed in your comments. I know it's early. Is that something that's going to be addressed directly?

DR. EMBRETSON: Well, you know, tests by their nature are involved in math proficiency and its promotion. We have the tests to gauge where the students are at and, you know, what we want to do to move them to another location.

But our concern is merely with validity in terms of what is being assessed and I think that's like a first question. What kind of proficiency will these tests promote exactly? That's the question of validity. So, it's being examined first and then given that we don't have a lot of time to put together a lot of research, we might not get as far as you want in that direction.

VICE CHAIRPERSON BENBOW: Bob?

DR. SIEGLER: Yes, I'd like to ask you to look up the road a little bit with regard to what you expect this state versus state and state versus NAEP set of comparisons to yield, because for sure there's going to be some overlap and they're going to be some differences. Neither the state tests nor NAEP are immune from criticisms. Many of them are the same criticisms, a few are unique to particular states or to NAEP. But there's no gold standard there. It isn't like anyone is just delighted with these tests, as far as I know.

So what you'll have, it seems to me, is this kind of set of descriptive results where it will be extremely difficult to make recommendations regardless of the particulars of how it comes out. Whereas, the international comparisons strike me as a more potentially promising way to go as well as relations to TIMMS, PISA and other international tests. I'm trying to think of what a good outcome of this set of comparisons would be.

DR. EMBRETSON: Well, I think a good outcome is to see how it interfaces with the other concerns from the other task groups. We can't do that until we know what the current situation is.

I mean, as far as the state level tests, we could have tests that are perhaps formally certain content, but they're tested at a very low level and they're tested with item formats that are not well regarded. We could look at that and we could make some conclusion about those kinds of items being not as represented in the definition of proficiency.

VICE CHAIRPERSON BENBOW: Valerie?

DR. REYNA: Thank you, Susan. And by the way, welcome to the Panel. I think you should take all this as a sign of tremendous enthusiasm for the topic that you're chairing.

I want to mention another theme that I think is beginning to emerge as a cross-cutting theme across the different subgroups, and that is this notion of computation skill versus conceptual understanding. We've seen that in the Learning Processes group and we saw that in that very informative summary that Deborah just gave in which she used those two subtopics to assess the effects of teacher training.

But I don't know whether that would fit under content, under cognitive complexity, and whether you can break that out. And what I'm talking about is, to what degree do tests assess computational skills, conceptual understanding, or both?

DR. EMBRETSON: Yes, that is intended under the complexity level factor for sure. Whether or not that can be reliably assessed looking at the items, that we do not know yet. So I would say that's unknown. I think we need to start with NAEP, which would be a smaller set and we need to look at to the extent to which we can classify that.

VICE CHAIRPERSON BENBOW: Sandra?

DR. STOTSKY: I'd like to just pick up on the point that was made that we have only had a few hours together and we're still in the process of clarifying some of the questions.

But one of the questions that we will try to make a little clearer will be, from at least my perspective, will be with regards to the state assessments. We'd like to look at how state assessments drive instruction, how they change instruction, if there's any research or literature available. Also we could look at how they drive teacher training and how they drive professional development, because they play key roles in all of these areas. So there is a lot more that can be put into this study beyond what its relationship to NAEP is. That is what we haven't had a chance to really think about i.e., how to develop more of these other areas that will be part of this one overall rubric.

VICE CHAIRPERSON BENBOW: Deborah and then Wilfried. Oh, Diane too. Wilfried, you're going to have to be last but you'll get your turn.

DR. LOEWENBERG BALL: Valerie's question

actually reminded me of something that I should have asked you a long time ago, which is, is this group planning to take up anything about teacher assessment?

DR. EMBRETSON: We had discussed that. It was on the list and at the moment it's not on the list. It probably is still somewhat open, I'd have to say.

DR. LOEWENBERG BALL: It might be good for at least the two groups to talk because obviously one thing that we're dealing with in our group is the psychometric quality of the measures being used to assess teachers' knowledge.

DR. EMBRETSON: Yes, well, it's a good idea, but at the moment we feel overwhelmed just looking at the student level. So, I don't know.

VICE CHAIRPERSON BENBOW: Diane?

DR. LOVELESS: Just a piggy-back, we actually, we were going to leave that to your group.

VICE CHAIRPERSON BENBOW: Diane?

DR. JONES: There's a lot of discussion among many of the groups about what is algebra, for example, what constitutes algebra. And I'm wondering if it would be a worthwhile experience for your group to maybe compare perhaps the eighth grade data with some of the much older tests. Maybe you could look at the tests that those of us in the post-Sputnik generation took, Iowa, Stanford, those are just the ones that I can think of. I wonder if it would be worthwhile to sort of compare what constituted competency then versus now to satisfy some of the questions about changes in our perception about what constitutes algebra.

I ask that because policymakers are who they are and many of them were engineers or people who came through school in the post-Sputnik generation and really, that's how we should be teaching algebra. It might answer some questions to sort of do a non-judgmental comparison about, you know, what were the standards then versus what are the standards now, and how might those standards differ in terms of computational ability and conceptual understanding.

DR. EMBRETSON: That is a good question in and of itself, I think. If we look at long term NAEP I'm not sure how well that represents it, but it does go back to 1978 at least. But you probably want to go back further.

VICE CHAIRPERSON BENBOW: Skip would like to address that and Skip is a member of the assessment group.

DR. FENNELL: Diane, only in the sense that the Conceptual Knowledge and Skills group is addressing what you said to an extent as we review curriculum frameworks and as we review textbooks and the like, but not so much in assessments, at least in this country.

So I guess my point is I don't know that we want to overload this group, but I think the point is well taken. We may have it covered at least to some extent in task group one.

DR. EMBRETSON: Well, if you have suggestions, the committee is certainly going to be open to hearing them. Some things that I have seen from other committees is the topic of patterns as algebra and that's something that we could look at. We also could look at estimation and its representation on tests. That is not a direct one right now. If you have some suggestions then we certainly want to hear them because I think that's the hardest area. If you're going to change the content you have to have more things being looked at in the items.

VICE CHAIRPERSON BENBOW: Joan?

DR. FERRINI-MUNDY: This is just quick and please don't interpret it as making more work for the assessment group.

Maybe it's tied to Val's question, but in looking at computational skill versus conceptual performance I'm wondering if something about real world problems might be included there as a piece of the complexity level discussion. Maybe our groups could connect on what we've seen and some definitions.

DR. EMBRETSON: Yes, I was just going to say, send us a definition.

VICE CHAIRPERSON BENBOW: All right, Wilfried has been very patient.

DR. SCHMID: Maybe I'm being presumptuous here, but it seems to me that the agenda that you outlined, which would include looking at various tests themselves and not just the outcomes, is just much more than you have time for.

I would suggest that you don't even attempt to look at test items or complexity, not because it's not important or interesting, simply because you do not have the time.

What I would suggest is that you take a very hard look at data that exists or are being generated about NAEP, various other tests and comparisons of NAEP to TIMMS and to other international tests. The question that Sandy raised, what is the effect of state tests practically speaking? Is it good? Is it questionable? I think that looking at the individual test items, test construction is just going to prevent you from doing something that you can't do in the time you have.

DR. EMBRETSON: Well, you may be right about that, because this committee is just starting now. Maybe we should have started earlier.

DR. LOVELESS: And if any committee members want to volunteer to serve on the Assessment Task Group we certainly would welcome you.

DR. SCHMID: I'll talk to you.

VICE CHAIRPERSON BENBOW: I'm on the committee and as we've all been working on committees we realize that we begin with big agendas. We're going to look at what's possible in the time frame that we have and we make compromises that sometimes hurt. We end up not being able to address issues that we would like to because we are time limited. Obviously, this is a very important question. We want to make sure we hit the right one and that we do prioritize our questions. Give us a little bit of a break. This could take a couple years to do well.

Well, at this point I'd like to bring this session to a close. I would like to thank the public and everyone for coming and attending and listening to us today. I would also like to announce that the next National Math Panel meeting will be hosted by Miami-Dade College in Miami, Florida on June 6th. So if you want to continue to hear the rest of the story, join us in Miami.

Thank you.

(Whereupon the meeting concluded at 12:11 p.m.)