U.S. DEPARTMENT OF EDUCATION

NATIONAL MATHEMATICS ADVISORY PANEL

Saturday, December 15, 2007

The meeting of the National Mathematics Advisory Panel convened in Salon ABCD, BWI Airport Marriott, 1743 West Nursery Road, Baltimore, Maryland, 21240 at 8:40 a.m.

PANEL AND EX OFFICIO MEMBERS PRESENT:

LARRY R. FAULKNER	Chair
CAMILLA PERSSON BENBOW	Vice Chair
DEBORAH LOEWENBERG BALL	Member
A. WADE BOYKIN	Member
DOUGLAS H. CLEMENTS	Member
SUSAN E. EMBRETSON	Member
FRANCIS (SKIP) FENNELL	Member
BERT FRISTEDT	Member
DAVID C. GEARY (Present by phone)	Member
RUSSELL M. GERSTEN	Member
TOM LOVELESS	Member
VALERIE F. REYNA	Member
ROBERT S. SIEGLER	Member
SANDRA STOTSKY	Member
VERN WILLIAMS	Member
IRMA ARISPE	Ex Officio
DANIEL (DAN) B. BERCH	Ex Officio
JOAN FERRINI-MUNDY	Ex Officio

PANEL AND EX OFFICIO MEMBERS NOT PRESENT:

LIPING MA	Member
JAMES H. SIMONS	Member
HUNG-HSI WU	Member
WILFRIED SCHMID	Member
RAY SIMON	Ex Officio
GROVER J. (RUSS) WHITEHURST	Ex Officio

U.S. DEPARTMENT OF EDUCATION STAFF PRESENT:

TYRRELL FLAWN	Executive Director
MS. MARIAN BANFIELD	
MS. HOLLY CLARK	
MS. IDA EBLINGER KELLEY	
MR. JIM YUN	

CALL TO ORDER

Chair Faulkner called to order the second day of 11th meeting of the Panel at Baltimore Washington International Airport. He indicated that they have signing services available, but they were not needed.

OPEN SESSION

NATIONAL MATHEMATICS ADVISORY PANEL: FINAL REPORT DISCUSSION

Chair Faulkner began with the principal messages section and the bullets related to Assessment, Instructional Practices, Teachers, Research, and Learning Process.

Dr. Gersten asked about item 9 and how the words tie into the standards that they adopted promising that early interventions are effective and promising. Dr. Siegler responded that they developed a set of rigorous standards that are not exactly the same as the Instructional Practices standards. The Institute for Educational Studies (IES) has endorsed several of these programs. Dr. Reyna added that, with respect to the major distinctions made in the standards of evidence document, these words are used consistently with that.

Dr. Gersten asked if there were corrections for clustering, because when those corrections are made, about 90% of studies reported in the literature as significant no longer are so. Dr. Reyna responded that the larger issue with clustering and other kinds of corrections has to do with things like replicability and so on. With respect to their report, they were in the fortunate position of having effects that in fact had been replicated over and over again, sometimes with diverse populations. The documentation is in the body of the report.

Dr. Schmid asked why "factual knowledge" crept in. Algebra material that might be thought of as factual knowledge really all comes under computational fluency. He thinks that the "factual knowledge" should be taken out. Dr. Sielger stated that in his community, computational fluency, mainly, is talking about algorithms beyond single-digit arithmetic. Factual knowledge is referring to single-digit arithmetic and is a fill-in for automatic knowledge of basic facts. Dr. Schmid stated that a big point is made of automaticity in the grid for pre-algebra topics. Dr. Reyna stated that "rapid retrieval" is perhaps not what they intend and therefore the word "knowledge" here does carry a little bit of the implication they want.

Chair Faulkner then moved to Learning Processes.

Dr. Fennell asked why they needed "accurately" there. Dr. Schmid stated that they should leave it there because there was a suggestion that the method of solution is more important than the accuracy of the answer, and that is something that they should not implicitly support. Dr. Benbow agreed because with gifted students, they often see that they can arrive at a solution but have no way of getting at the solution. They want them to accurately solve the problem, not just achieve the correct answer.

Dr. Stotsky asked about the section that says, "U.S. children do not reach the point of fast and efficient solving of single digit…" The next paragraph indicates that textbooks in the U.S. present easier single-digit arithmetic problems far more frequently

than harder single-digit problems. She does not see the two points logically relating to each other. Dr. Siegler responded that the relation is that in comparisons with Singapore, within the set of single-digit problems, U.S. workbooks present problems like 2 + 2 more often than problems like 7 + 9. There also is another problem with fluency with more complex algorithms. Students have problems with both.

Dr. Schmid stated that he has not done a formal study, but he thinks may be true of single-digit arithmetic and multi-digit arithmetic. He is troubled with singling out the single-digit arithmetic problems, and would prefer to omit single-digit. Dr. Geary responded that he is okay with the single-digit issue. The point is, to get equal fluency for easier and harder problems, the harder problems need to be practiced more frequently than the easier problems. Dr. Berch asked if Dr. Stotsky's point was that there seemed to be a conflict that the first statement didn't specify the difficulties in computational fluency as being with harder single-digit problems. And then they talk about how such an emphasis on easy problems is why they would expect students to be recalling those in a fluent fashion.

Dr. Siegler added that one way to address Dr. Stotsky's point would be to insert the term "relatively difficult" in paragraph three before single-digit addition. It would say, "many contemporary U.S. children do not reach the point of fast and efficient solving of relatively difficult single-digit problems, and much less fluent execution of more complex algorithms. Dr. Geary stated that there was a study about 20 years ago that reviewed textbooks in the United States, Japan, Russia, and so forth, and it was exactly as Dr. Schmid said.

Dr. Schmid does not agree with the contradiction that Dr. Stotsky sees. If they are talking about efficient performance of single-digit addition, if students practice that all the time and don't address more difficult single-digit, students will not be proficient in single-digit addition. He would like to keep what is there, with an extension of the point that problems in the U.S. textbooks are too easy, from just single-digit arithmetic to whole number arithmetic.

Chair Faulkner asked if there was agreement that there is a deletion of "single-digit" right before "arithmetic" in paragraph four and after "harder." It would say, "In the United States, easier arithmetic problems are presented far more frequently than harder problems." Dr. Fennell asked if they know this and Dr. Siegler responded that they did because of a study by Karen Fuson. Dr. Fennell stated that this was 20 years ago. Dr. Geary stated that he did a study 10 years ago.

Chair Faulkner called a vote on the motion to delete "single-digit" in two places in paragraph four, which was agreed upon.

Chair Faulkner then moved to item 6.

Dr. Schmid stated that there are two different occurrences of "accurately," and this reference is to "accurately estimate" and he has a problem with that. Dr. Fennell agreed that the issue here is estimation and the level of approximation that one gets. They need the word accurately there. He also asked about "debilitating mathematics anxiety," and if they want to say that. Dr. Siegler responded that he agrees on the debilitating point, but not at all on the accurate estimation. A problem with the little estimation instruction that goes on is that there is a large emphasis on using varied strategies, which is good only if the estimate is accurate.

Dr. Siegler added that in his research, he found that college students at Carnegie Mellon thought it was more than plausible that it's a good estimate to say that the population of Norway is approximately 40 million people. However, it's an inaccurate estimate. The actual population of Norway is about 7^{th} or an 8^{th} of that. So there is a difference between "accurate" and "plausible." "Accurate" has a rather straightforward meaning and it's the right term to use.

Dr. Schmid stated that when one estimates, there is such a notion as estimating within the certain order of accuracy. The adverb "accurately" is inappropriate. Chair Faulkner stated that "appropriate" and "appropriately" have been accepted elsewhere in this text. The use of "appropriately" was approved.

Dr. Siegler stated that "debilitating" is too extreme a term. It was not meant to draw too much concern about students who feel a little bit of anxiety, as most do. They might say "serious." Dr. Schmid asked if they could drop "debilitating." They want to reduce anxiety whether it is debilitating or not. Dr. Reyna stated that some anxiety is good. Dr. Boykin stated that one of the oldest laws in psychology, the Yerkes-Dodson effect, does suggest that there can be an optimum level of arousal that can actually facilitate performance, and to the extent that those were unchangeable, then there is a level at which it might not be problematic. Dr. Fennell stated that "serious" is fine.

Chair Faulkner then moved to Instructional Practices.

Dr. Fristedt asked about formative assessment being based on the state standards, and whether they are also using sample items from state assessments. The Assessment Task Group found that assessments don't always match standards. Dr. Gersten replied that they are from the state standards not from the state test. Dr. Siegler stated that there is a change in perspective that weakens the strong point that could be made here. It recommends regular use of formative assessments for students in the elementary grades and what teachers should do. The second sentence says that these assessments need to provide information not only in their content validity, but also in their reliability and their criterion related validity. It's not talking about teachers, but is talking to researchers. It might not rise to the level of importance for the executive summary. Dr. Gersten stated while they wanted more information on the psychometric characteristics, it belongs in the body if the text, as opposed to the executive summary. Dr. Berch stated that the sentence needs more explanation.

Chair Faulkner asked if there was agreement that the middle sentence in the second paragraph comes out, and there was.

Dr. Clements asked to take out the added text. Dr. Gersten asked to delete the final sentence and delete that extra verbiage. Dr. Siegler agreed that the last sentence shouldn't be there. He suggested staying with the first sentence, and then saying, "Teacher's regular use of formative assessment improves the student's learning, especially if teachers have additional guidance on using the assessment to design to individualized instruction. Although the research to date has only involved one type of formative assessment…" He then suggested saying, "the results are sufficiently promising that the Panel recommends regular use of formative assessment for students in the elementary grades." It includes the qualifier, but it isn't as downbeat.

Dr. Gersten agreed with that. Dr. Schmid asked if the new phrasing is now only one type of formative assessment without the explanatory note. He would leave in,

"based on items sampled from the major curriculum objectives for that year." Dr. Siegler agreed with that.

Chair Faulkner moved to item 23, "high quality studies show that a particular cooperative learning strategy…."

Dr. Fristedt asked if this point could be clearer about the strategies used. Dr. Loveless stated that this is a cooperative learning technique that has been around for 25 years, and it's well known in the field. It is described in detail in the body. Dr. Gersten agreed that it is known in the literature, but there is a problem that it is not a term used anymore for a variety of reasons. It's part of one of Success for All, which is used in the lot of schools. A sentence in the summary that explains the essence could be useful. Cooperative learning is simply breaking the class into three to six groups and doing something. Dr. Loveless stated that there is language from the body that they can use for this.

Dr. Siegler stated that the inclusion of the peer-assisted instruction along with the Team Assisted Individualization (TAI) is confusing here. Given that its benefits are so qualified, he recommended taking that sentence out. Dr. Loveless agreed that they should take the peer-assisted sentence out. The nature of the evidence supporting the two is quite different. They only have two studies of peer-assisted learning that had significant effects and the same research team conducted them.

Dr. Gersten stated that some mention of peer-assisted learning should be in there. Though they are only the two studies, they were very large-scale studies compared to the smaller ones on TAI. The effects were well replicated. In addition, people are familiar with the term peer-assisted learning. It is currently used in the country and it is always nice to recommend things that people can get their hands on, as opposed to things from the historical archives.

Dr. Loveless disagreed and said they would have to put those two pools of evidence side-by-side to compare them. It was only the classroom-level effects that were significant, and not the student-level effects. Dr. Ferrini-Mundy agreed. Dr. Gersten stated that it's a pretty arcane statistical point, but in terms of the principles, Dr. Reyna is talking about replication. That is not a good way to convert one to the other. Dr. Ferrini-Mundy stated that she is not opposed to taking it out, but if they do, they are essentially endorsing one specific and commercial style of cooperative learning.

Dr. Fristedt stated that when he was speaking about the specificity, he had something different in mind. Dr. Stotsky asked whether they were necessarily endorsing something by pointing out what the finding is. She wondered if the beginning of the sentence should state, "high-quality studies on a variety of cooperative learning strategies show that one particular one…." It's really the details of this one particular type that are important and most people don't know that there is a whole world of them out there.

Dr. Berch asked if they could describe the TAI without using that label. This would acknowledge that although there are a number of cooperative learning strategies, only one has shown it to be effective for computational fluency. Dr. Loveless added that this is a strategy that is a proper noun. They could add to the executive summary some language that would provide that kind of detail.

Dr. Boykin asked what the justification was for putting this in here since it is new language. He also asked why TAI is a proper name while peer-assisted instruction is something generic. Chair Faulkner stated that it might not be clear that they are talking

about two different approaches. Dr. Schmid recommended a footnote, which refers to a later discussion of what actually characterizes this proper noun, TAI.

Chair Faulkner then moved to item 24, real-world problems.

Dr. Schmid stated that he has a problem with the statement that there are high-quality studies that show that if teachers teach X, then students are better at X. Dr. Ferrini-Mundy stated that what's interesting about the findings on real-world problems is that these problems are often included in curriculum materials with the intention that it will enable students to learn what they might call straightforward mathematics. The finding is that there isn't any demonstrable impact on mathematics that might be tested through computations, operations, or basic applications. The impact is on near and far transfer of being able to solve real-world problems better.

Mr. Williams asked if they could remove the second sentence as the first sentence says everything. Dr. Reyna agreed, but if the intervention were-- to improve X, improve X-- they wouldn't be sitting here. The fact that evidence shows that something aimed at achieving a certain goal did so, that is important news. Dr. Fennell agreed with Dr. Reyna that they need the second sentence describing the importance of application of topics.

Dr. Ferrini-Mundy stated that she would like to just use the first sentence, and then a little bit of editing to pickup some of what is in sentence two. Dr. Schmid stated that it should say that real-world problem solving is a valuable skill and there is evidence that teaching it actually helps to achieve it. He is bothered by the use of "stronger than that of students who receive traditional instruction," when traditional instruction is not defined. He would rephrase the second sentence. Dr. Siegler stated that he is concerned about how the general public will read this point without understanding the limitations. They need to be careful of what they are claiming. Dr. Fristedt added that the phrase "real world" is a loaded term. Mr. Williams offered to work on this point.

Dr. Clements asked to move 23 behind 21, which was agreed.

Chair Faulkner then moved to item 25 on Explicit Instruction for Students.

Mr. Williams asked if there should be a description in parenthesis about implicit instruction. Dr. Sielger added that it could say explicit instruction. Dr. Gersten stated that it means that students have many opportunities to ask and answer questions, and think aloud about decisions they made. Dr. Loveless added that it is defined at the end of the previous paragraph in the "i.e." Dr. Siegler added that it means students are provided with opportunities to think aloud, and talk through the decisions they make and the steps they take. Dr. Gersten agreed.

Dr. Fennell asked if they were saying "daily" or if "regularly" might be more appropriate. This research was based on a study involving children with learning disabilities. He asked if they are being overly prescriptive for such students. Dr. Gersten replied that in terms of the set of studies, there were approximately 18 to 20 high-quality studies with students with learning disabilities, which typically means they're about in the lower 10^{th} percentile. There also was another smaller set of about 8 studies, which was with low-achieving students, the lowest third on national tests. They pooled those together with some trepidation. He agreed on the substitution of "daily" for "regularly."

Dr. Schmid stated that the wording of, "This finding does not mean that all students' mathematics instruction should be delivered in an explicit fashion," could be read as meaning, don't do it for all students. What they really want to say is that the research was directed at the low-performing students. Chair Faulkner stated that he read it

to mean that it's not necessary that all the instruction, even for low performing students, be delivered in an explicit manner.

Chair Faulkner asked a group to work on the rewording of that, and stated that they would take off "daily," and insert "regularly."

Dr. Fristedt asked if in item 25, the term "word problems" is used as a synonym for real-word problems as used in 24, or if it means something different. In addition, where it says, "i.e., instruction where teachers clearly model approaches for solving problems…" he likes that description, but wonders if it supported by the research. The word "model" is fine, except it is a word that also has a life of its own in curriculum matters. He wonders if "illustrate" would be better than "model." Dr. Gersten stated that model is a word he believes teachers are used to hearing. Dr. Stotsky stated that the general public might not commonly understand it.

Dr. Siegler stated that “some” is better than “regularly.” Dr. Berch asked about the order, which doesn’t seem to be consistent with what's in the draft. Chair Faulkner stated that this happens in several areas due to the logic of what they are trying to deliver in a summarized form.

Dr. Siegler stated that item 25 should be brought to a shorter conclusion.

Chair Faulkner moved to item 26, "Engaging Students with High Quality Software."

Dr. Schmid asked if this was consistent with the study that Dr. Whitehurst brought up. Dr. Clements replied that it was not consistent with other studies, but Dr. Whitehurst’s study is well represented as one of the studies in the tutorials and drill and practice meta-analysis. That is in the body of the report.

Chair Faulkner moved to item 27, “Based on the review of 12 studies that meet the Panel's rigorous criteria, calculators have shown a limited to no impact."

Mr. Williams stated that this is misleading, as the details of how the calculators were used in the studies are not there. Dr. Schmid stated that the calculators were used in a very limited way in the studies. Dr. Fristedt stated that one has to focus on the specific way the calculator is used, because this is such an emotional issue. Dr. Stotsky recommended that they right away refer to the National Survey of Algebra Teachers as the recommendation in place of what is there for 27.

Dr. Schmid stated that the proper place to discuss the detail of the studies is in the Learning Processes report. Additional detail would be taken as meaning that calculator use does no harm, which is not a message supported by evidence.

Chair Faulkner asked if there was agreement to take the first recommendation on page 49 and use it to replace 27.

Dr. Fennell stated that item 27 is not a recommendation, but a finding. It’s important to say that the research in this field is not only limited, it's also old. In addition, calculator use with young children hasn't been adequately investigated. They owe it to the field to say that there is a tremendous need to research the impact of calculator use. Dr. Siegler stated that they could leave the first two sentences as they are on 27, because that is what was found. Then they could mention that the Nation's Algebra Teachers Survey indicated that the use of calculators in prior grades was a concern of teachers. Following that would be that research is needed to indicate the effects of long-term use. Mr. Williams still disagreed.

Dr. Schmid stated that the studies do not show that calculators have shown limited to no effect in calculational skills, et cetera. What they do show is that when calculators are used, the way they were in those studies, then there is limited to no effect. So they definitely do not show that calculators have limited to no impact on computational skills. Dr. Clements added that each study used calculators differently. They are old studies, but the re-analyses revealed two marginally significant effects. The vaguely positive empirical evidence doesn't get any play in these kinds of descriptions, but the huge negative effects receive a lot of attention.

Dr. Geary asked why they would say that calculators do no harm if interventions with calculators are not showing that student are doing better. Dr. Schmid stated that he was concerned with how they were stating the findings of these 12 studies where calculators have limited to no impact. Dr. Reyna suggested noting that there is some intermediate level evidence that is relevant here. There is good evidence that fluid retrieval of arithmetic facts or automaticity is related to ability with fractions and obviously with other concepts. There is good evidence that automaticity is achieved through repeated practice. She would add, “to the degree that calculator use supplants opportunities to practice retrieval, it will interfere with mathematical performance because automaticity is not achieved.” If it decreases the opportunity for practice, that's one thing; if it increases the opportunity represent information in a better way, that’s another.

Dr. Fristedt stated that “calculators have been shown” needs to be changed to “particular uses of calculators have been shown,” which is a different issue.

Chair Faulkner called a vote on elevating the recommendation on page 49 into the list of points in the executive summary. The second vote was whether to eliminate or to modify the paragraph relating to scientific findings.

Dr. Ball would like to bundle the recommendations together, particularly the ones about teachers, and then the recommendations relating to the Learning Processes group would be together. Dr. Fennell agreed they need a clarion call for the field about the role of the calculator in teaching and learning mathematics. They have to ask for support of research relative to the use of calculators. Dr. Schmid suggested keeping the recommendation as stated on page 49 with two modifications: The first sentence goes and the last sentence is replaced by "To the degree that calculator use supplants opportunities to practice retrieval, it will interfere with mathematical performance, because automaticity is not achieved."

Dr. Ball stated that adding the last sentence seems peculiar, because the recommendation now stresses the need to study the effects of calculators and then concludes with a claim about the effects of calculator use. It seems logically inconsistent to first say, they need to study this, and then say, but they know this other thing. Dr. Reyna stated that the last sentence is based on sets of data on the fluent retrieval of arithmetic affects--facts affecting other kinds of mathematical performance. There is also data that automaticity is achieved of repetitive practice. The third sentence is about the degree that calculator use supplants opportunities to practice. Dr. Ball stated that they should state that as a separate point that it's an analysis of other things they know and doesn't go with the recommendation for research.

Dr. Loveless stated that he would like to keep the item on Algebra teachers.

They have a body research that doesn't detect the problem, but practitioners think there is a problem. They have to look at both of those.

Dr. Schmid stated that they would rephrase sentence one, keep sentence two, and then finish with Dr. Reyna's suggestion. They would also find a place for the Algebra teacher finding.

Dr. Boykin stated that it is important to note that in these 12 studies, they did cover methods that looked at a variety of patterns and extended use of calculators.

Chair Faulkner stated that the motion is to duplicate the recommendation of page 49 in the executive summary, which was agreed upon.

Dr. Ball moved that they separate what they just moved from 49 into three separate blocks: one is the report of the survey, two is the phrase "in light of…pursued," and the third is something about what the Panel recommends. Dr. Siegler stated that it might not be a good idea to devote four points to calculators, which is probably the least agreed on and most contentious issue before the Panel.

Dr. Loveless stated that they want to acknowledge that the body of research is limited in some ways, but it does have a rather benign finding when it comes to calculator use. Then a second paragraph could just start with, "However, algebra teachers say that students use calculators too much. Also, they know that automaticity is very important and to the degree the calculators interfere with that, it would obviously interfere with math learning."

Chair Faulkner stated that a group would work on that item

Chair Faulkner then moved to item 28, "Mathematically Gifted Students with Sufficient Motivation." He stated that Dr. Geary had comments about social and emotional functioning.

Dr. Benbow stated that the studies they included in the Task Group report did not directly address social emotional development. Other studies address this, and there is no impact on social emotional development. Dr. Benbow stated that she would like to add the words after learning, "and should be allowed to do so," because this is a finding and not a recommendation. Mr. Williams asked if they could insert, "challenging mathematics," as well.

Dr. Fristedt asked if they could add the phrase, "at a much higher rate and in much greater depth." Secondly, has asked to separate out the recommendation of making arrangements for students. He asked if they also could say something about grouping since these students learn so much from each other. Dr. Benbow stated that they really did not look at that issue in their research.

Chair Faulker stated that the motion was to take the, "and should be allowed to do so" and convert it into a second paragraph, which was not agreed to. The next motion was whether the Panel accepts "and should be allowed to do so."

Dr. Schmid recommended that they leave out "successfully."

Chair Faulkner stated that "should be able to learn challenging mathematics successfully at a much higher rate and in a much greater depth" is probably overkill. The group agreed that "successfully" would come out. Dr. Ferrini-Mundy asked if the research says anything about the "the greater depth" or if it's more about the pace. Dr. Benbow stated that to be really consistent with what the research said, it was much more about the rate and the pace than about depth. If they say "challenging mathematics," that encompasses depth. Dr. Schmid stated that he had no problem with the intent but

somehow if they say "challenging mathematics at a higher rate," they are combining things that really can't be combined. He asked if they are saying that all students should learn more challenging mathematics or that the accelerated students should learn more challenging mathematics. Chair Faulkner stated that "challenging" and in "much greater depth" shouldn't both be there. Dr. Siegler suggested they proceed with "and should be allowed to do so," with the deletion of "successfully."

Dr. Berch stated that there are three points in there. First, it's not saying anything about the students other than the mathematically gifted. Second, it states that mathematically gifted students can learn at a faster pace even if they don't have "the challenging" content. He asked if they are also sure that they can learn even more challenging math at a faster pace.

Chair Faulkner called for a vote on the original language, with "successfully" removed, and "then should be allowed to do so" added. That was approved.

Chair Faulkner then moved to item 28, "Curricular Content."

Dr. Fristedt asked who the audience was for items 1 through 4, and Chair Faulkner said the general public. Dr. Fristedt stated that the target audience should be book publishers. Dr. Schmid stated that they should not change 1 through 4, but they might be able to elaborate later on who this addresses.

Chair Faulkner asked about the use of the word multi-digit. Dr. Fennell stated that he would like to keep the original wording, which is, "By the end of Grade 3, students should be proficient with the addition and subtraction of whole numbers. At the end of Grade 5, students should be proficient with multiplication and division of whole numbers." The insertion of digit type makes these benchmarks too targeted and minimalist. Dr. Schmid stated that putting in multi-digit in some sense weakens it because what they are talking about is a variety of aspects of these operations which are independent of whether it is one digit or several. The Benchmarks for the Critical Foundations and the Critical Foundations for Algebra spell out in detail what robust sense of number is, including automatic recall and fluency with the standard algorithms.

Chair Faulkner confirmed that they should leave out multi-digit.

Dr. Sielger stated that it is independent how quickly one goes through what sometimes are three separate Algebra courses and whether it's presented in an integrated or segregated way, so they might not have the basis for saying this. They have made the point prior that all students, regardless of which course they go into, need this to be able to take advanced math their high school senior year. Dr. Schmid stated that the phrasing could be changed. What the paragraph says is the way integrated curricula and state standards are set up at the moment means that students taking those courses don't get to take calculus in high school.

Dr. Stotsky suggested this be stated as, "As they're currently implemented." Dr. Siegler asked if they knew this from the research. Dr. Stotsky stated that this is based on the research of the Institute for Defense Analyses/ Science and Technology Policy Institute (IDA/STPI). They looked at the differences between a state's standards or objectives for these three courses, for the single subject course sequence, and the state standards offered for the curricular sequence, and found this great disparity. Dr. Fennell stated that this was just for one state, and Dr. Stotsky stated that is was for North Carolina.

Chair Faulkner stated that the motion is to approve everything except the last paragraph, which was agreed upon.

Chair Faulkner then moved to the executive summary on assessment, item 32.

Dr. Stotsky recommended deletion of the second sentence. Dr. Fristedt agreed that it didn't add anything. Dr. Embretson stated that the phrase "should focus" should change to "adequately represent," because focus can mean a lot of things and they are talking about representation in their materials. Dr. Fristedt stated that "adequately represents" better describes the situation as there are other foci for the assessments, as well. Dr. Schmid suggested they leave the first sentence and omit the second sentence, and in the third sentence change "focus" to "adequately represent." Dr. Siegler stated that putting "adequately represent" in the first sentence would be weird, because it implies that they also should include information that students shouldn't learn. Dr. Loveless stated that he would keep the word "focus" throughout because it is a stronger term than "adequately represent." It doesn't mean that the test is confined to this material. It just simply says this is the emphasis of the test, but this should be the focus.

Dr. Boykin stated that the fact that it hasn't been spelled out more specifically suggests that someone has a concern about the test design process taking these factors into consideration. There is some concern that these different players are not involved in test design or test construction. Dr. Embretson stated that she was fine with not having "adequately represented" in the first sentence. It can be put in the third sentence because what test developers understand is representation in the form of a blueprint. Either they refer to it in terms of test design and blueprints, or put "adequately represent" somewhere to take care of that concern.

Dr. Loveless stated that it would be easy to construe that the status quo adequately represents it. Dr. Schmid replied that he sees it the opposite way, if they are saying it should adequately represent. If they say "focus," that seems much more consistent with maintaining that currently everything is okay. Dr. Embretson stated that that would be the case since they then address the Critical Foundations of Algebra.

Dr. Berch stated that he was uncomfortable with the first statement. He asked if they could combine those to say that the mathematics that students should learn should guide the design of National Assessment of Educational Progress and state tests. Dr. Fennell stated that it seems like the main point of the discussion is the use of the words "focus on." He stated that they should begin the statement very directly and would argue for keeping the language as is, but they could delete the second sentence.

Dr. Embretson stated that without some idea of representing a different balance, the second sentence has to stay something about the design to represent content.

Dr. Benbow made a motion on the third sentence, "For Grades 4 through 8, the tests should focus on and adequately represent the Panel's Critical Foundations." They would also remove the second sentence.

Dr. Sielger stated that to avoid people asking why they should not be represented on state tests in Grades 1, 2, and 3, they should say something like, "On the NAEP in Grades in 4 and 8, and on state tests through Grade 8, the tests should focus on and adequately represent…"

Chair Faulkner called a vote on, "NAEP and state tests for students through Grade 8 should focus on and adequately represent the Panel's Critical Foundations of Algebra.

Student achievement on this critical mathematics content and should be reported and tracked over time." This was agreed upon.

Chair Faulkner then move to item 33.

Dr. Embretson recommended they move 34 to follow 32. Dr. Stotsky asked about mentioning the balance of algebra and patterns. Dr. Fennell stated that it doesn't rise to the level of the executive summary. Dr. Siegler stated that this is an extremely important item and they don't want to dilute the focus on the two strands. Dr. Schmid agreed that because there is tremendous overemphasis on the pattern problems in the current NAEP and state assessments that it should be addressed in the executive summary. Dr. Fennell stated that this point addresses looking at number differently. Dr. Benbow also thinks that the algebra point should be a separate recommendation that follows. They could say "rebalance and with less emphasis on patterns."

Chair Faulkner stated that there was a question about decimals. Dr. Sielger stated that his concern about fractions is that performance is low even on tasks such as comparisons. Because performance is so poor, tests will not discriminate between students who know more and less. Dr. Embretson stated that because it is a matrix-sampling kind of design and not individual scores, they can have some fractions and decimals on the test and it will not degrade anything else. Dr. Sielger withdrew his objection.

Dr. Fristedt stated that something about a number line representation of fractions and decimals is appropriate. Dr. Benbow stated that she is not sure it rises to the level of the executive summary.

Dr. Benbow moved that the paragraph say, "The Panel recommends a more appropriate balance in how algebra is defined and assessed for both the fourth- and eighth-grade levels of NAEP. The Panel strongly recommends that algebra problems involving patterns should be greatly reduced in NAEP and state tests."

Dr. Loveless stated that there is problem with that because state tests are not matrix tests. The state tests are given to every student. So they need to treat NAEP and state tests a little differently in that regard. Dr. Schmid asked if they could have a separate sentence for the state assessment, something to the effect that the same considerations apply to the state assessments.

Chair Faulkner called for a vote on that motion to create a new item 34, which was agreed up. He then moved to item 33, which would now be 35, that state tests and NAEP must be of the highest technical and mathematical quality.

Dr. Embretson stated that the statement doesn't seem to apply to item development. The problem lies much deeper than needing to have higher levels of expertise involved in item development. She asked to add, "To this end, states and NAEP should develop better procedures for item development, quality control, and oversight to ensure…" Dr. Stotsky added that they should say, "reflect the best item design features." Dr. Benbow suggested replacing "design principles" for "features."

Dr. Embretson stated that there are all kinds of lists of practices and it really hasn't gone far enough. She is talking about design features that are based on research

Chair Faulkner called a vote on that language change, which was agreed upon. He then moved to item 36 on calculator use on assessment.

Dr. Embretson stated that items are usually constructed by item quality developers to measure one kind of thing so they don't put in all kinds of other sources of difficulty. If

they let calculators be used, they include complicated numbers in the real-world problems, which isn't the goal. That introduces a source of difficulty because some people may misuse the calculators. Dr. Schmid stated that calculator use hinders automaticity, but they have this recommendation that is a reflection of this fear at the assessment level. They want to test certain things that cannot be tested adequately with calculators. Dr. Fristedt stated that some assessment items should test more than one thing. They stayed away both in Conceptual Knowledge and Skills and in Assessment from talking about constructive uses of the calculator and he does not think they can get into that here.

Chair Faulkner called for a vote on the motion, and it was agreed upon.

Dr. Siegler asked if they should add a sentence right after the one they just approved to say, “This recommendation should not be taken to preclude inclusion of items designed to assess effectiveness of calculator use.” They stated this because it is something they want students to be able to do when appropriate. Dr. Schmid stated that in no way does the sentence suggests that there shouldn't be calculator use on the tests. What it speaks to is simply the fact that calculators are too broadly used on assessments. Dr. Boykin stated that the statement is very strong. Dr. Embretson added that she does not want to add something that encourages more calculator use than what is already out there.

Chair Faulkner moved to item 37 on the “continuum of research.”

Dr. Ferrini-Mundy stated that the sentence should read, “Both smaller scale experiments on the basic science of learning and larger scale randomized experiments examining effective classroom practices are needed to ensure the coherent growth of research addressing important questions in mathematics education. Basic research is as rigorous as randomized trials, and is necessary to develop explicit predictions and to test hypothesis which are under emphasized in current research on mathematics education.” Mr. Williams asked if they could make this a separate recommendation that deals more with the need for future research.

Chair Faulkner called for a vote on that motion, which was agreed upon.

Chair Faulkner moved to math specialists, and Instructional Materials point 29. The second sentence, which now starts out, “excessive length makes our book unnecessarily expensive,” should be edited. It should be changed to, “excessive length makes our books more expensive.”

Chair Faulkner called for a vote on that motion, and it was approved.

Chair Faulkner then moved to item 27, which is the revised calculator item.

Dr. Loveless read the new statement, which was, “A review of 12 studies that met the Panel's revised criteria, only one study less than 20 years old, finds limited and no impact of calculators on calculation skills, problem solving, or conceptual development. The finding is limited to the effect of calculators as used in the studies. However, the Panel’s survey of the nation's algebra teachers indicated that the use of calculators in prior grades was one of their concerns, to the degree that calculators impede the development of automaticity, fluency in computation will be adversely affected. The Panel recommends that high-quality research involving both short and long-term effects of calculator use on computation problem solving and concepts be pursued.”

Dr. Fristedt stated that he does not like the phrase about calculators not having an impact. It is only in the next sentence that they find out that it is a very restrictive use that

is being talked about. He would like to modify it to say, “limited to no impact of calculators as used in these studies.” He also would like the Panel to recommend high-quality research involving both short and long-term effects of calculators on various uses of calculators.

Dr. Clements stated that they could insert that phrase. He took Dr. Loveless’ sentence about high-quality research as indicating that, because in the full report they criticize the research for not giving them adequate descriptions of how calculators were used. They could say “high-quality research involving both short and long-term effects of particular uses of the calculator.” Dr. Fennell added, “both short, long-term, and particular effects of calculator use on computation problems and concepts…”

Chair Faulkner confirmed that the change would say, “A review of 12 studies that meet the Panel's rigorous criteria, only one study less than 20 years old, found limited to no impact of calculators on calculation skills problem solving or conceptual development. These findings are limited to the effect of calculators as used in the 12 studies. However the Panel's survey of the nation's algebra teachers indicated that the use of calculators in prior grades was one of their concerns, to the degree that calculators impede the development of automaticity, fluency and computation will be adversely affected. The Panel recommends that high-quality research on particular uses of calculators be pursued, including both their long and short-term effects on computation problem solving.” He called a vote on that motion and it was agreed upon.

Dr. Berch addressed the Learning as we Go Along section in the principle messages. He moved that they delete that paragraph. Chair Faulkner called a vote on that and it was approved. He then moved to the edits to item 24.

Dr. Ferrini-Mundy read the edits, which were “A small number of high-quality studies indicates that if mathematical ideas are introduced using real-world context, then students' performance on assessments involving real-world problems is improved, although there is no significant difference on their performance on calculation and routine procedures.” Dr. Berch stated that taking out the comparison group takes away the need to say, “significant difference,” so they just have to say there is no impact or enhancement.

Dr. Fristedt stated that they should change, “a term not well defined in the literature,” to “a term that is used in quite different senses at various places in the literature.” Dr. Stotsky suggested they say, “A small number of high-quality studies indicate if mathematical ideas are introduced using real-world context…” and then say that it is defined in specific ways or very carefully defined. Dr. Ferrini-Mundy stated that the way it has been reworded is a pretty accurate description. It is a set of studies, which when they pull the effects, even though they use slightly different definitions and different approaches, they find a significant effect size. It would nice to be able to say more but they use these words in different ways. Performance when assessments include real-world problems is significantly stronger when instruction emphasizes real-world contexts. And there isn’t a difference on performance on the parts of the assessments that are about calculations or routine procedures. Therefore, it impacts near and far-term transfer.

Dr. Stotsky stated that when they talk about calculation and routine procedures, it is not clear what is intended. Dr. Ferrini-Mundy stated that they avoided the traditional instruction issue by not bringing it up. She stated that Dr. Stotsky was talking about

outcome measures. In that area, they do have information about what was measured at the end.

Dr. Loveless stated that the previous version of this began with a limitation and that was done on purpose because of the fact that real world is often sold as a panacea. That version stated that, "use of real-world problems does not enhance learning mathematics as measured by assessments not calling for application." And then he would add the rest of Dr. Ferrini-Mundy's proposal after that.

Dr. Stotsky added that they need to say, "use of real-world problems, a term that is not clearly defined and researched, does not enhance learning mathematics as measured by assessments not calling for application." Dr. Berch stated that some of the message gets lost because they are trying to fit it all into one sentence. Dr. Schmid added that there should be an introductory sentence.

Dr. Benbow asked a group to work on that issue. She then moved on to item 23 on cooperative learning approaches.

Dr. Loveless read the revised version, which was, "research on a variety of cooperative learning approaches shows that students' computational skills are improved by one cooperative learning strategy, TAI. This strategy involves heterogeneous groups of students helping each other, individualized problems based on students' performance on a diagnostic test, and awards based on both group and individual performance. Effects of TAI on conceptual understanding and problem solving were not significant. Research on peer-assisted learning indicates the potential of this approach for improving students' computational skills. This is based on two studies and additional research is needed. The two studies were confined to Grades 2, 3, and 4."

Dr. Boykin stated that it should be clear that peer-assisted learning is also considered as one of the types of cooperative learning approaches. Dr. Loveless stated that cooperative learning involves groups of four or five students. In TAI, the peer-assisted learning groups were just two students. Mr. Williams asked if all TAI strategies involve heterogeneous grouping or if some involve just whatever the groups might be. Dr. Loveless replied that he would have to go back and look at those studies, but he is pretty sure they all involved heterogeneous grouping. Dr. Boykin added that the basis of heterogeneity is ability level. Dr. Berch stated that it needs rewording because it says that research on all of these shows that only one has effects. Dr. Loveless stated that he could drop the peer-assisted mention because it has a pretty modest evidentiary base that then triggers these reader questions.

Dr. Fennell suggested that they keep the paragraph relative to TAI and refer to that as the single cooperative learning strategy where they found evidence. The rest of this, which is the other strategy of peer-assisted learning, would appear later on in the report. He moved to accept the first paragraph.

Dr. Benbow called a vote on the motion and it was agreed to take it out.

Dr. Berch stated that the first sentence on TAI should be modified. It is not the research on all those approaches that shows that this particular approach is the one. Mr. Williams stated that "among a variety of approaches this one was shown…" Dr. Loveless suggested, "research has been conducted on a variety of cooperative learning approaches. One cooperative learning strategy, TAI, has been shown to improve students' computation skills."

Chair Faulkner moved to item 16, "using value-added analysis…"

Dr. Boykin stated that he was stuck on the notion of greater gains and asked, “greater than what?” Dr. Loveless suggested that they say “significant.”

Chair Faulkner then moved to item 17.

Dr. Stotsky suggested they change the word “supports” to “confirms,” because it makes the research follow what common sense would suggest. She also asked about the purpose of measuring and using value-added measures. Dr. Ball stated that they wished they could say exactly what teachers should be taught. The can't do that, so they need measures that permit them to understand exactly what aspects of mathematical knowledge have a bearing on teachers’ instructional skill and their students' learning.

Dr. Berch recommended switching “shows the strongest signal” for “provide the strongest indicator.”

Dr. Fennell asked that for parallel structure. If they talk about elementary and middle school first, then towards the end it should also say elementary and middle.

Dr. Benbow then moved to item 18.

Dr. Stotsky asked if they could change it to, “coherent and focused induction and professional development.” This would be clearer that they are talking about preparation programs and alternative pathways. Dr. Benbow stated that that is not what the Task Group found. They reviewed research in all these different aspects of teachers' education, the four types of teachers' education.

Dr. Ferrini-Mundy stated that the last sentence says, “Extant evidence suggests that there are not significant differences among current pathways,” and asked if this referred literally to the nature of those pathways or about their outcomes. Dr. Ball stated that they are different from each other.

Dr. Benbow moved to number 18.

Dr. Stotsky stated that this item is preceded by an opinion, and other items are not.

Dr. Benbow moved to item 21.

Dr. Loveless stated that “supporting” isn’t right because the salary schemes don't support teachers' effectives. They are based on teachers' effectiveness. Dr. Benbow suggested stating, “for teachers' effectiveness.”

Dr. Benbow moved to math specialists.

Dr. Ball stated that math specialist includes three forms that they reviewed. They are lead teacher model, math coaches, and fulltime teachers. They can mention that parenthetically. Dr. Stotsky agreed that math coaches should be there, and asked if they could also add pullout teachers. Dr. Ball stated that they did not review pullout teachers.

Dr. Schmid stated that the message was that nothing in the existing research strongly supports the use of either math specialists or math coaches. That does not come across as clearly as it did in the original phrasing. Dr. Ball stated that the group rejected that yesterday and that’s what this was a response to. There was a heavy argument yesterday that said, while they know that mathematical knowledge of teachers is related to their student's achievement, they don’t know if the model of a full-time elementary teacher would work.

Dr. Schmid stated that there is currently an army of so-called math coaches. Dr. Fennell stated that this point on research on the elementary math teachers' specialist is offered due to the past discussion. They could make a parenthetical reference to "all of the models," and then go into a discussion that specifically targets the elementary math

teacher model. Mr. Schmid stated that the fact that there was little support for math coaches does not come across. There should be a distinction in the language between a math specialist who deals with students and math specialists who deal with teachers only.

Dr. Ball stated that they do not have research on either model. Dr. Loveless stated that it sounded as if they were advocating something on which there was no research base. But it did not have anything to do with the inclusion of math coaches later on in the bullet.

Dr. Schmid stated that he would like to say that they don't have research justification for the use of either model. However, they have justification for different reasons, namely practicality and maximizing the amount of teacher knowledge in mathematics. Dr. Loveless stated that he continues to object to that because they do not have any evaluations of that kind of policy or its consequences. Dr. Berch stated that that is why they developed this paragraph regarding going beyond the practicality and the need to explore whether it's really helpful.

Dr. Loveless suggested they say, "However, they found high-quality research showing that the use of these specialist teachers, including math coaches, improves students learning." Dr. Ball stated that they agreed to put a parenthesis in with the names of the three models they reviewed.

Dr. Benbow then moved to the item in Instructional Practices on, "make use of what is known from rigorous research about instructional practices."

Dr. Loveless stated that he would like to lead with what is currently the last sentence. Even though they have a laundry list of practices that research does have some encouraging things to say about, they still don't add up to very much. They should lead with, "teachers' own professional judgment and experience should prevail on the areas where research is not available…" and then add the things that have been shown to be effective in research.

Dr. Berch added that there should be a qualification about master teachers or teachers who have been shown to be effective. Not every teacher can do it. Mr. Williams added that he would like it to say, "Because they found no evidence to support the superiority of any particular instructional practice, classroom teachers led by master classroom teachers who have produced academically successful students should be allowed and encouraged to make instructional decisions based on their students' academic background and their classroom experience." Dr. Benbow added that they did find some effects with TAI. Mr. Williams stated that this would give teachers some power to make decisions based on what they know works.

Dr. Schmid stated that New York City has strict rules about the number of minutes of direct instruction allowed. There are strict rules about how to arrange the furniture in the classroom. This is the sort of thing to which they object. Dr. Ferrini-Mundy asked if they could start with, "master teachers' professional judgment and experience should prevail in areas where research is not available."

Dr. Stotsky suggested they not start with "where research is not available" because there will always be somebody who will find a study or claim that there is research. Dr. Ferrini-Mundy stated that they should say something about a body of rigorous research that gives robust findings. Dr. Berch stated that they could say, "Where high-quality research exists, teachers should make use of those to guide their instructional

practices." Where that does not exist then, they should talk about the importance of their professional judgment.

Mr. Williams stated that the sentence about the teachers making a choice should be first so they can say something to teachers. They have found little to say to teachers, but the one thing they did find was that there was no evidence to support the superiority of a particular instructional practice.

Dr. Schmid stated that they should be consistent with how they are supposed to arrive at recommendations. Then they can strongly criticize straitjackets on teachers.

Dr. Loveless stated that they should lead with the affirmation of teacher's professional judgment and experience, and begin by recognizing that. Secondly, he would make the point that where high-quality research exists, not naming all these different things, those findings should form teachers' practice.

Dr. Berch stated that he wouldn't argue that if it were coming out of a different Panel and a different report. They need to be guided by the terms of the Instructional Practices report. Dr. Fennell stated that there should be a statement that acknowledges the need for research to inform practice, and that when there is not an evidentiary base, a well-intentioned master teacher has the opportunity and should be encouraged to make the kinds of decisions that they're professionally capable of making.

Dr. Benbow called a vote to see if the research comes first, which was agreed upon.

Chair Faulkner then asked the Panel to pass a motion accepting the task group reports and transmit them into the production process

Dr. Benbow stated that the task group reports are in the name of the task group members only, not as members of the whole Panel. So all they are doing is accepting a report. They are not approving those reports.

Chair Faulkner stated that the minutes of this Panel need to have an entry that this Panel did receive the reports and that they were preceding with production.

Dr. Schmid stated that he first would like to see the final language on calculators and the rest of the Instructional Practices report. Dr. Loveless also stated that he didn't want to vote to receive the Instructional Practices report because he had not read the latest version that came out this week.

Dr. Schmid moved for the Conceptual Knowledge and Skills report to be received. Dr. Benbow seconded that motion. The vote was unanimous.

Chair Faulkner moved for the Instructional Practices report to be received. The vote had one opposed.

Dr. Berch moved for the Teachers Report to be received. Nine were in favor, two opposed, and Mr. Williams abstained, along with two others.

The Assessment report was moved with one abstention.

Dr. Ball moved that the Instructional Practices report be received. Dr. Ferrini-Mundy seconded that. Dr. Loveless stated that he would vote against it.
Chair Faulkner asked if they should table that until the next meeting. Dr. Schmid agreed. Dr. Fennell asked the chair of that Task Group to respond to postponing it. Dr. Ferrini-Mundy stated that there would be more work on the report as everybody knows, but the notion of accepting and receiving the report seems innocent enough. She would like it moved but she recognizes some of her colleagues are not going to vote in favor of this.

Dr. Boykin asked if it would still be sent into production. Chair Faulkner stated that it would after it has been finished, but the minutes will show that the Panel didn't receive it if they don't receive it.

Dr. Ferrini-Mundy asked what “not accepting” it would mean. Did it mean that people are expecting that exact language from what they talked about today to be in the report? Chair Faulkner stated that he did not believe that would be necessary. What would be necessary is for what they have done in the executive summary to be completely reflected in the body of the final report. Chair Faulkner asked if they wanted to wait until January or February to receive this.

Chair Faulkner called the vote to receive the Instructional Practices report. Five accepted and eight opposed.

Chair Faulkner then called for a vote on the motion to receive the reports of the Subcommittees on Instructional Materials and the National Survey of Algebra Teachers, and that was agreed upon.

The session adjourned at 3:19 p.m.

I certify the accuracy of these minutes.

Chair Signature_____Date_____

Vice Chair Signature_____Date_____

A**DDENDUM: PUBLIC PARTICIPANTS**

First Name	**Last Name**	**Organization**
Judy Ann	Brown	Words and Numbers, Inc.
Elizabeth	Carson	NYC HOLD National
Cathie	Dillender	Pearson
G. Stanley	Doore	East County Citizens Advisory Board Member
Alice	Gill	American Federation of Teachers
Cheryl	Jaffe	Northrop Grumman Electronic Systems
Henry S.	Kepner, Jr.	University of Wisconsin-Milwaukee
Dr. Genevieve M.	Knight	Knight and Associates
Ken	Krehbiel	National Council of Teachers of Mathematics
Lisa	Lavelle	DC FAME, Maret School
Michele M.M.	Mazzocco, Ph.D.	Johns Hopkins University
Leah Casey	Quinn, Ph.D.	Montgomery County Public Schools
Judith	Reed, Ph.D.	National Council of Teachers of Mathematics
Robert A.	Richards	Sadlier-Oxford, division of William Sadlier, Inc.
Melanie A.	Ryan, MSEd.	JUST MATH
Jacqueline	Smith	National Council of Teachers of Mathematics
Larry	Snowhite	Houghton Mifflin Company
Gerald E.	Sroufe	American Educational Research Association
Janie	Zimmer	National Council of Supervisors of Mathematics