

SUMMARY OF REPORT  
to  
NYS BOARD OF REGENTS  
and  
COMMISSIONER OF EDUCATION

Presented by the Math A Panel

on

October 8, 2003

# INTRODUCTION

THANK YOU!

The Panel deeply appreciates the invitation to assist with this important effort.

GRATEFUL FOR HELP.

The Panel is grateful for all of the help from SED staff members.

THIS IS A CONSENSUS REPORT.

All 41 recommendations are supported by all 13 Panel members.

# INTRODUCTION, cont'd

The Panel believes that, because of the length and complexity of this report, we need to be very clear in stating what we intend the report to say.

# WHAT THE REPORT DOES *NOT* SAY:

- This report does NOT say there are some problems with the Math A testing program, so throw out Math A.
- The report does NOT say there are problems with the Math A testing program, so don't give kids math Regents exams any more.

Conclusions like these would be a misinterpretation of this report.

# THE REPORT *DOES* SAY:

- There are problems:
  - with the clarity of the Math A standards
  - with the support systems needed for such an ambitious endeavor, and
  - with some of the technical aspects of the development of the Math A Regents exams.
- These problems are solvable; 41 recommendations were developed to address the problems identified.

# THE PANEL BELIEVES . . .

- NYS has the highest math standards in the nation; we want to keep it that way.
- Much positive change has occurred with the new math standards, in particular access for hundreds of thousands of kids to high quality programs that were not available to their counterparts a decade ago.
- We must move forward, not backward.

# THE PANEL BELIEVES . . .

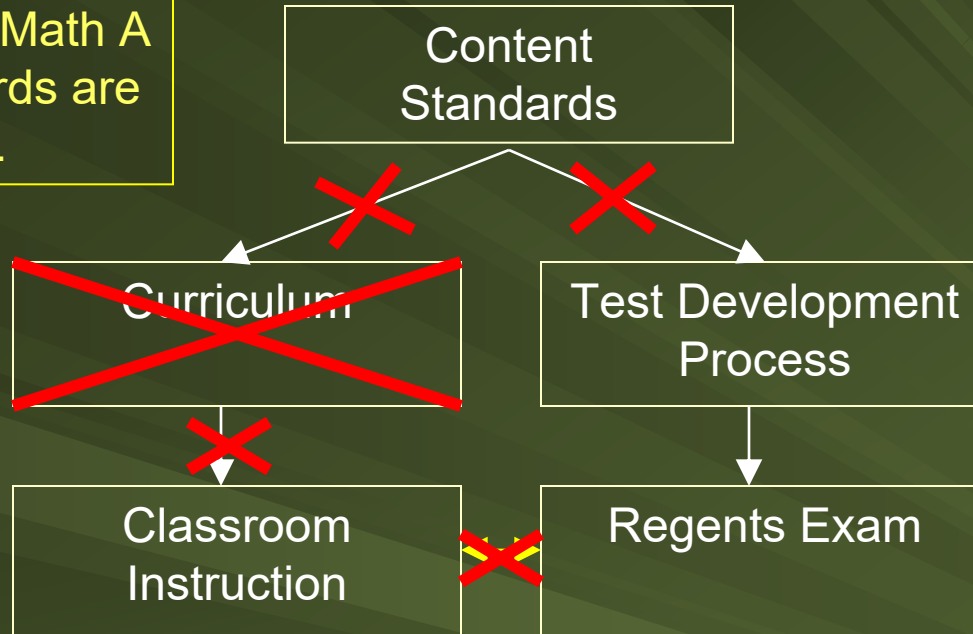
- The problems with Math A are complex and systemic.
- Therefore, the solutions, of necessity, are complex, and must be systemic.
- The recommendations should be considered as a whole. Accepting even a majority of the recommendations may not result in any change. We believe that all the pieces need to be in place for this plan to work.

# ALIGNMENT IS ESSENTIAL

FINDING: The Math A content standards are not clear.

FINDING: There is no agreed upon curriculum.

FINDING: Because the standards are unclear, and there is no agreed upon curriculum, teachers look to *past Regents exams* for guidance.

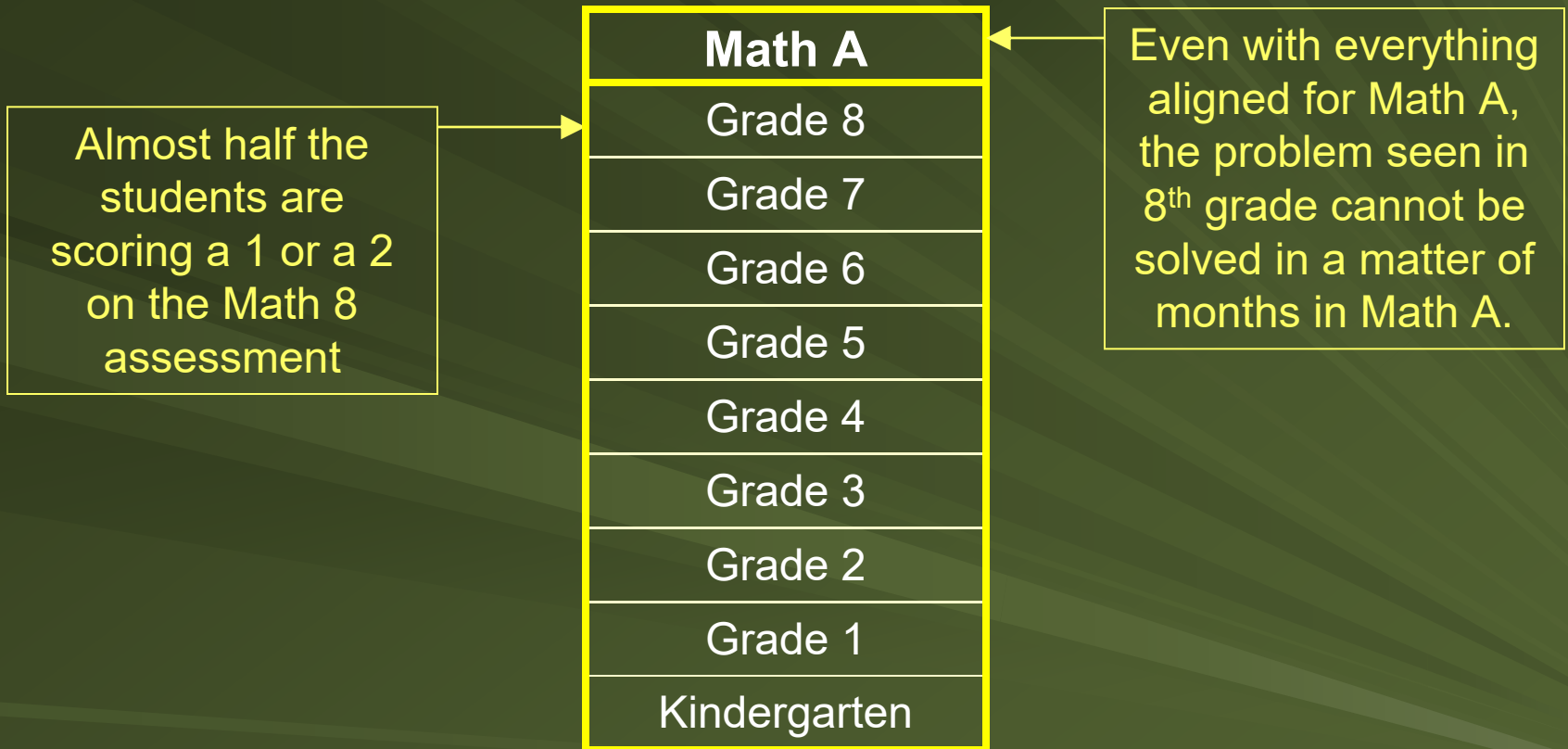


FINDING: There are technical problems in the test development process, which result in the exams being *inconsistent* over time.

**FINDING: THE SYSTEM IS NOT ALIGNED.**



# MATH A IS BUILT ON A K-8 MATH FOUNDATION



The effort to raise all students to the Math A standards **must start in the early grades.**

# ITEM RESPONSE THEORY

**Classical Test Theory.** Each Item is given a “p” value estimating the probability of getting the item correct. This is simply the percent of students in the field sample that get it correct. If 75% of the sample get it correct, then  $p(\text{item}) = 0.75$ . p-values range from 0 to 1.

# ITEM RESPONSE THEORY, cont'd

**Item Response Theory.** Uses mathematical formulas to transform the p-value to another number, called the Rasch Item Difficulty, commonly called a “b-value.”

- The range is usually -3.0 to +3.0. More negative means the item is relatively easy; more positive means it is relatively harder; 0 is “average.”
- These can be thought of as roughly similar to standard deviations:
  - most items will fall between -1.0 and +1.0;
  - vast majority between -2.0 and +2.0;
  - -2.0 very easy, +2.0 very hard;
  - -3.0 anyone can get; +3.0 no one can get.

# MATH A STANDARD SETTING

At the first round of testing (98-99), the items were arrayed by their b-values, from easiest to hardest. A large group of people met and spent time deciding where the line should be drawn for passing (65) and passing with distinction (85). The lines drawn were:

- 0.692 for Passing (65)
- 1.612 for Passing with Distinction (85)

# STANDARDS – A CLARIFICATION

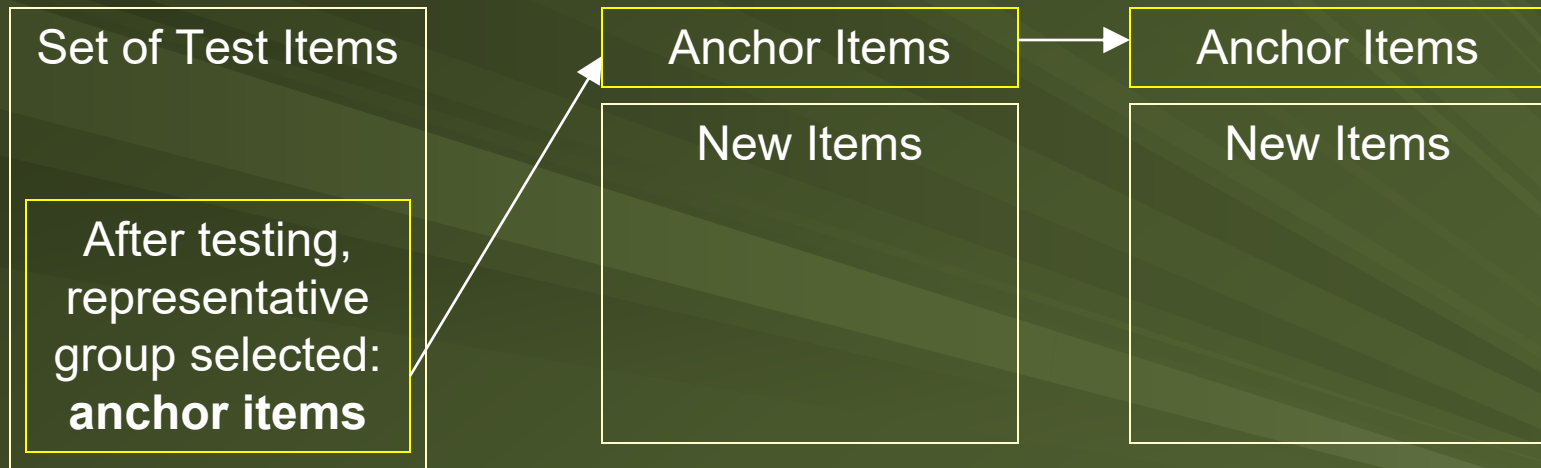
- **“Content Standards”** are statements of expectations about student knowledge and skill in the content area. (For example: “Students will be able to solve algebraically two simultaneous equations with two variables.”)
- **“Performance Standards”** are the set performance levels for the exam. (For example: 0.692 for Passing – 65 -- and 1.612 for Passing with Distinction – 85.)

# THE IMPORTANCE OF ANCHOR ITEMS AND FIELD TESTING

TEST  
DEVELOPMENT  
CYCLE #1

FIELD TEST  
TEST  
DEVELOPMENT  
CYCLE #2

FIELD TEST  
TEST  
DEVELOPMENT  
CYCLE #3



In each test development cycle, the estimated difficulty of each new item is determined based on the field test group's performance on the anchor items.

# IMPORTANCE OF FIELD TESTING

**The performance of the field test students on the anchor items is the basis for adjusting the b-values of all newly written items.**

(For example, if the field test students score lower than the first test group on the anchor items, the conclusion is that the field test group has weaker skills. Therefore, their weaker skill level is used to adjust the b-values of the new items accordingly. An item the field test group experiences as difficult will have its b-value adjusted, using mathematical formulas, to correct for the skill level of the group.)

**The purpose of IRT is to establish item difficulty *independent from* the skill level of the group.**

# IRT HINGES ON:

- The anchor items, how well they represent the content standards, and how well they “behave” statistically over time.
- The size of the field test group. (NOTE: Not its representativeness, because the mathematical adjustment of the item difficulty should correct for skill level.)



# RAW SCORE TO SCALE SCORE CONVERSIONS

The conversions each year are all based on the standards determined in 1998-1999: 0.692 for Passing (65) and 1.612 for Passing with Distinction (85).

Scale Score	Raw Score			
	June 2002	August 2002	January 2003	June 2003
65	52	53	52	51
85	72	72	71	70

**THE MATH A PANEL'S  
FINDINGS AND  
RECOMENDATIONS**

# SIMPLY PUT: THE STANDARDS ARE NOT CLEAR

- The anchor items look like they represent one standard.
- Many of the Regents exam items look like they represent a different standard.
- The examples in the guides to teachers look more like the anchor items.
- Many Standards (Performance Indicators) are worded so broadly that their intent is unclear; they could be interpreted in many ways.

# FINDINGS AND RECOMENDATIONS

## THE MATH A STANDARDS

- **Finding 1:** *The Math A standards lack clarity and specificity (p. 15).*
- **Recommendation 1A:** *Educationally useful standards must be developed in mathematics for each grade, K-8, and for Math A and Math B, that consist of a clear, well-defined set of skills, the mastery of which is demonstrable (p. 19).*
- **Recommendation 1B:** *SED should establish a mathematics standards committee to rewrite the standards into functional form, and to meet regularly in the future to analyze test results, thus ensuring continuous relevance (p. 19).*
- **Recommendation 1C:** *SED should develop and disseminate suggested curricula for mathematics instruction for each grade K-8, and for Math A and Math B (p. 19).*
- **Recommendation 1D:** *To benefit from the extensive research and deliberation of the current Math A Panel, some of the current Panel members should be included in both new committees recommended in this report, i.e., the standards committee, and the curriculum development committee (p. 20).*

# FINDINGS AND RECOMENDATIONS THE MATH A STANDARDS

- **Finding 2:** *The design concept that the Math A exam should be taken by the typical student after three semesters of instruction has not been successful (p. 20).*
- **Recommendation 2:** *The standards and curricula should be structured so that the typical student will take the Math A exam after one year of high school mathematics (p. 20).*

# FINDINGS AND RECOMMENDATIONS

## THE MATH A EXAM

- **Finding 3:** *The June 2003 Regents Math A exam was harder than past Math A exams (p. 25).*
- **Recommendation 3A:** *Establish and maintain narrow statistical targets for difficulty of Parts I, II, III, and IV of the Math A exam forms (p. 28).*
- **Recommendation 3B:** *Review the Math A item pool (p. 28).*
- **Recommendation 3C:** *The difficulty of problems in the anchor item set, in the guidance documents provided to teachers, and on the actual tests should be aligned (p. 28).*
- **Recommendation 3D:** *The weighting of the open-ended items, number of scale points possible on the open-ended item rubrics, and other aspects of the scoring of open-ended items should be reconsidered (p. 29).*
- **Recommendation 3E:** *The Math A test should focus on a more limited, more clearly-specified set of content standards and indicators (p. 29).*

# FINDINGS AND RECOMMENDATIONS

## THE MATH A EXAM

- **Finding 4:** *The Math A tests have not been able to maintain a consistent performance standard over time (p. 29).*
- **Recommendation 4A:** *Alternative equating designs should be considered (p. 32).*
- **Recommendation 4B:** *Sampling procedures for estimating item performance must be improved (p. 32).*
- **Recommendation 4C:** *Replace the anchor item set (p. 33).*
- **Recommendation 4D:** *Revisit performance standards (cut scores) (p. 33).*

# FINDINGS AND RECOMMENDATIONS

## THE MATH A EXAM

- **Finding 5:** *The New York State Education Department cannot accurately predict performance on Math A test (p. 33).*
- **Recommendation 5A:** *SED should implement procedures for predicting the performance of test forms and groups of students on future Math A exams (p. 34).*
- **Recommendation 5B:** *Policies for field testing and data collection should be revised (p. 34).*



# FINDINGS AND RECOMMENDATIONS

## THE MATH A EXAM

- **Finding 6:** *Support and oversight for the Math A exam program should be improved (p. 34).*
- **Recommendation 6A:** *SED should immediately increase in-house content and technical expertise resources by a minimum of one psychometrician and two math content specialists (p. 35).*
- **Recommendation 6B:** *SED should clarify the responsibilities assigned to its technical advisory committee, and should request this group to provide regular reports, including technical analyses, reactions to proposed changes in test programs, and suggestions for improving State testing programs (p. 35).*
- **Recommendation 6C:** *SED should increase demands placed on contractors (p. 36).*
- **Recommendation 6D:** *Internal coordination and documentation should be improved (p. 36).*

# FINDINGS AND RECOMMENDATIONS

## THE INFRASTRUCTURE

- **Finding 7:** *Passing rate data for the State as a whole were not available until three months after the exam; no data are collected regarding student performance on individual items, nor even regarding student performance on the four parts of the exam (p. 37).*
- **Recommendation 7:** *SED should increase its data collection capacity to include item level data, and should accelerate its data collection timetable (p. 37).*
- **Finding 8:** *While the most important use of student performance data is to inform instruction, statewide data mining models that would enable local schools and teachers to use these data effectively are not generally available (p. 37).*
- **Recommendation 8:** *SED should substantially broaden its efforts to assist districts in data collection, and the use of data to inform instruction (p. 37).*

# FINDINGS AND RECOMMENDATIONS

## THE INFRASTRUCTURE

- **Finding 9:** *The mathematical background of teachers delivering math instruction varies widely; yet, raising almost three million children to higher levels of math achievement will be impossible without highly skilled teachers (p. 37).*
- **Recommendation 9A:** *SED and higher education need to continue and to strengthen their partnerships to ensure strong teacher education programs, both pre-service and in-service (p. 37).*
- **Recommendation 9B:** *The certification requirements for elementary teachers and special education teachers should include a minimum of nine credits of college level mathematics (see Recommendation 9C), and three credits of teaching techniques in mathematics (p. 37).*
- **Recommendation 9C:** *Mathematics courses required for certification, both for mathematics teachers and elementary and special education teachers, should be specific not only in terms of number of credits required to be taken, but also in terms of coursework required to be taken, e.g., calculus, number theory, algebraic structures, probability and statistics, etc. (p. 38).*

# FINDINGS AND RECOMMENDATIONS

## THE INFRASTRUCTURE

- **Recommendation 9D:** *The Panel believes that, for any teacher responsible for teaching mathematics at any level, the 175-hour professional development requirement should include specific mathematics requirements. The Panel's thinking is that:*
- *teachers who teach mathematics exclusively should be required to take 100 of the 175 hours in the area of mathematics;*
  - *secondary teachers who are certified in, and who teach in, more than one subject area, should be required to take 50 of the 175 hours in the area of mathematics;*
  - *teachers who teach mathematics as part of a broad set of teaching responsibilities, e.g., elementary teachers and special education teachers, should be required to take 30 of the 175 hours in the area of mathematics.*
  - *Additionally, the range of possible courses that would satisfy these requirements should be clearly specified (p. 38).*

# FINDINGS AND RECOMMENDATIONS

## THE INFRASTRUCTURE

- **Finding 10.** *The public has very little awareness of Math A, and may have misunderstandings about the goals of Math A (p. 38).*
- **Recommendation 10:** *Make greater use of SED communications capacity to engage the public in conversations about the importance of strong mathematics skills (p. 38).*
- **Finding 11:** *There is often a "disconnect" between K-12 and higher education (p. 38).*
- **Recommendation 11:** *SED should encourage conversations at the local and regional levels of K-12 teachers of mathematics and higher education professors of mathematics, for the purpose of sharing curriculum, and exploring professional development opportunities and other possible collaborations, to bridge the gap between K-12 and higher education (p. 38).*

# FINDINGS AND RECOMMENDATIONS

## THE INFRASTRUCTURE

- **Finding 12:** *Raising the level of mathematics achievement of all students to high levels must start when children are very young, and must go beyond the school day for school aged children (p. 39).*
- **Recommendation 12:** *SED should encourage through grants and other means the expansion of mathematics education initiatives beyond K-12, such as the creation of partnerships between schools and libraries, and the greater use of public television and museums (p. 39).*

# FINDINGS AND RECOMMENDATIONS

## ADDITIONAL ISSUES --

### SCORING RUBRICS, AND COMMUNICATION TO THE FIELD REGARDING GRADING

- **Finding 13:** *The scoring rubrics do not give credit for a variety of mathematically correct approaches (p. 40).*
- **Recommendation 13A:** *Develop more generally worded, holistic scoring rubrics which permit credit to be granted for atypical, but mathematically correct, student responses (p. 40).*
- **Recommendation 13B:** *Rubrics should be designed so students do not lose 33% or 50% credit for a minor arithmetic error (p. 40).*
- **Finding 14:** *There is a serious "disconnect" between the perception of the SED content specialists and the perception of field classroom teachers regarding the application of the scoring rubrics (p. 40).*
- **Recommendation 14:** *On each set of directions for the Math A exam, a statement should be added confirming that the scoring rubrics are a guide and should be applied using professional judgment (p. 40).*

# FINDINGS AND RECOMMENDATIONS

## ADDITIONAL ISSUES --

### SCORING RUBRICS, AND COMMUNICATION TO THE FIELD REGARDING GRADING

- **Finding 15:** *There needs to be better communication of SED grading interpretations during the grading process for the Math A exams (p. 41).*
- **Recommendation 15A:** *SED should continue on its path of setting up a website during Math A Regents exam grading to provide up-to-date clarifications to teachers grading the exam (p. 41).*
- **Recommendation 15B:** *SED should explore ways of sending up-to-date grading clarifications to the school districts during the grading period following the administration of the exam, as a backup to the website, to ensure the greatest possible consistency of grading across the State (p. 41).*



# FINDINGS AND RECOMMENDATIONS

## ADDITIONAL ISSUES --

### CALCULATOR USE ON THE MATH A EXAM

- **Finding 16:** *Allowing the option of using a graphing calculator on the Math A exam provides some students with an advantage on the exam, thus creating an inequitable situation (p. 41).*
- **Recommendation 16:** *The use of calculators on the Math A Regents exam should be standardized (p. 42).*

# RECOMMENDATIONS

## CONCERNING THE JANUARY 2004 EXAM, AND ALL MATH A EXAMS UNTIL A NEW ONE IS DESIGNED

- **Recommendation 17:** *Until the standards are rewritten, new curricula are developed, the new course is delivered, and a new Math A Regents is designed and field tested, the Math A Regents exam should be restructured so the exam includes: 30 Part I items, 5 Part II items, 2 Part III items, and 2 Part IV items (p. 43).*
- **Recommendation 18:** *The exam should be reviewed by a group of practitioners, including math teachers, university mathematicians and mathematics educators, with representatives from this Panel, prior to the administration of the exam (p. 43).*
- **Recommendation 19:** *Until new items are developed and properly field tested, the exam items should be scaled in accord with the procedures used for the August rescaling of the June 2003 exam (p. 43).*

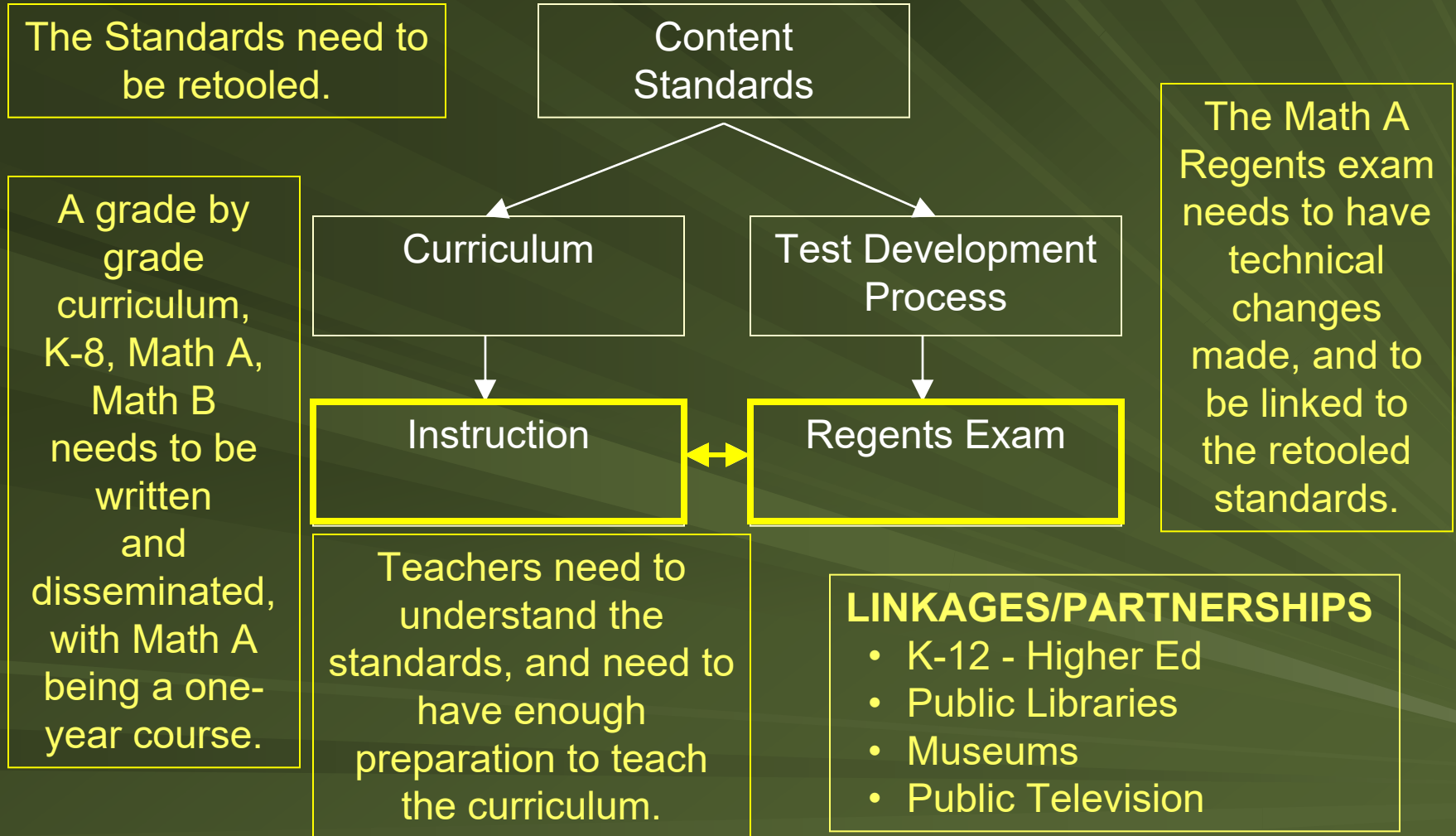
# RECOMMENDATIONS CONCERNING THE JANUARY 2004 EXAM, AND ALL MATH A EXAMS UNTIL A NEW ONE IS DESIGNED

- **Recommendation 20:** *The scaling should not be finalized until after the exam has been administered and after a post equating procedure has been implemented to ensure the fairness of the test (p. 43).*
- **Recommendation 21:** *The 55 passing option on the Math A Regents Exam for a local diploma should be continued until after the standards have been clarified, after new curriculum has been developed and disseminated, and after a new exam has been developed and administered for at least one school year (to ensure that it is performing in accord with its design) (p. 44).*
- **Recommendation 22:** *The math RCT safety net for special education children should be continued until after the standards have been clarified, after new curriculum has been developed and disseminated, and after a new exam has been developed and administered for at least one school year (to ensure that it is performing in accord with its design) (p. 44).*

# AN OFFER OF HELP

The Panel believes this work is important. Panel members have offered to continue helping by having representatives serve on the standards committee, on the curriculum committee, on any certification/professional development committees, and on the review committees for future exams -- especially the January 2004 exam -- and recommends this continued involvement to the Board.

# SUMMARY



# CONCLUSION

➤ The Math A Panel believes these 41 recommendations:

- will align all the elements of this effort
- will greatly strengthen Math A
- will result in a strong K - Math B program
- will improve mathematics education for all of the children in New York State

➤ We hope our thoughts are helpful and that they will, ultimately, serve those for whom we are all here, the children.