Smarter Balanced Assessment

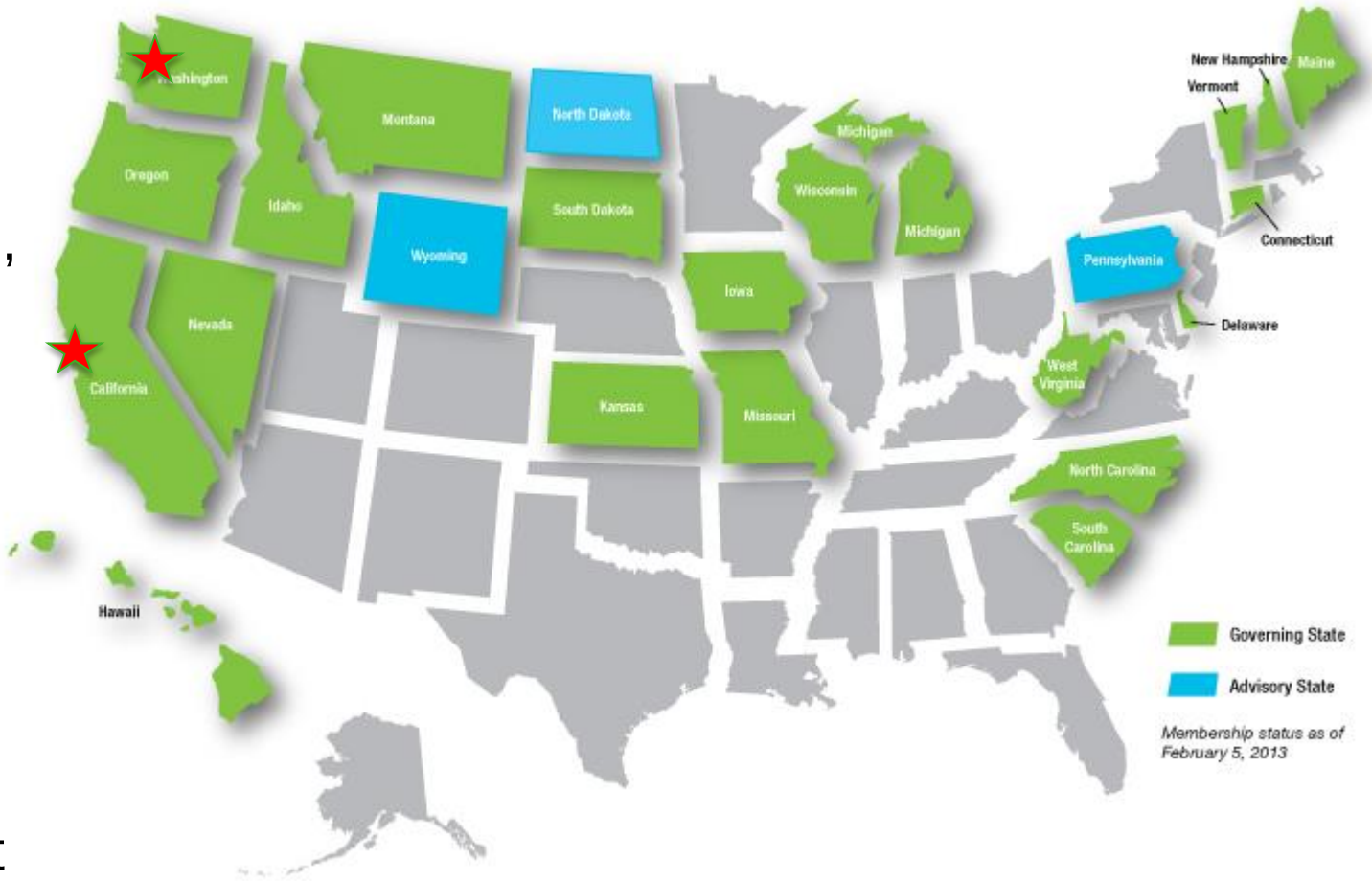
Developments and Challenges

Joe Willhoft & Shelbi Cole
MSRI Conference
April 5, 2013

Smarter Balanced Assessment Consortium

A National Consortium of States

- 24 states representing 39% of K-12 students
- 21 governing, 3 advisory states
- Washington state is fiscal agent
- WestEd provides project management services



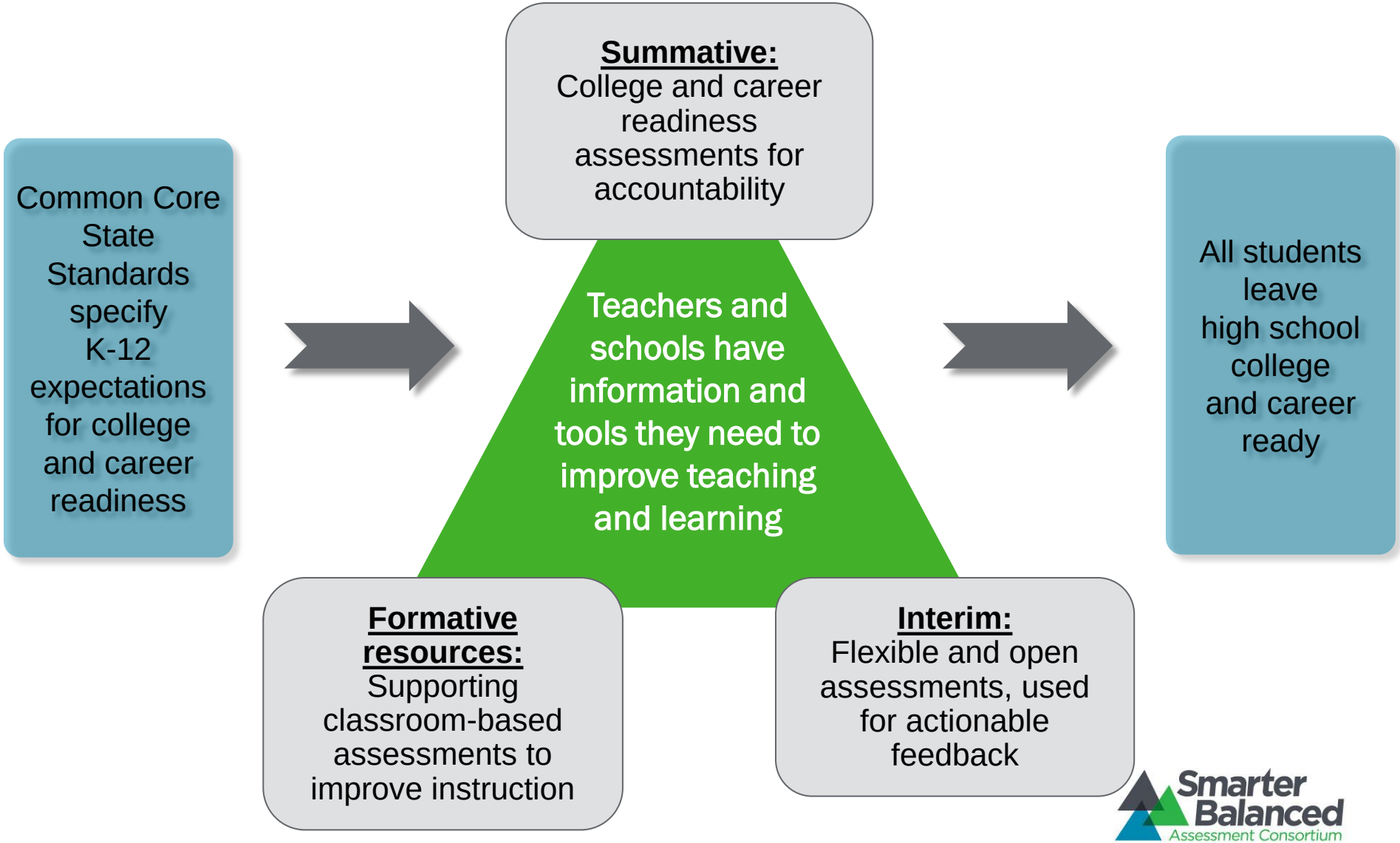
Smarter Balanced Assessment Consortium

A Next Generation of Assessments

- Alignment to Common Core State Standards in Mathematics and English language arts/Literacy
- Rigorous assessments showing progress toward "college and career readiness"
- Common cut scores across all Consortium states
- Providing achievement and growth information
- Valid, reliable, and fair for all students (except those with "significant cognitive disabilities")
- Using multiple measures
- Administered online
- Operational in 2014-15 school year

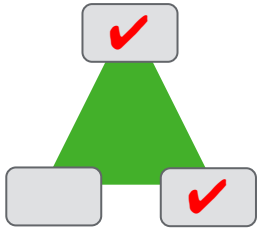
Smarter Balanced Assessment Consortium

A Balanced Assessment System



Smarter Balanced Assessment Consortium

Computer Adaptive Technology



Faster results, fewer items	• Turnaround time is significantly reduced • Can assess broad range with fewer items
Deeper Analysis	• Classroom and school reporting captures the full range of items seen by many students
Increased precision	• Accurate measurement across range of students • Improved measures of student growth over time
Tailored to student ability	• Item difficulty based on student responses
Greater security	• Large item pool means not all students receive the same questions
Mature technology	• GMAT, GRE, COMPASS (ACT), Measures of Academic Progress (MAP)

Smarter Balanced Assessment Consortium

Score Reports for Mathematics

Claim #1 - Concepts & Procedures	"Students can explain and apply mathematical concepts and interpret and carry out mathematical procedures with precision and fluency."
Claim #2 - Problem Solving	"Students can solve a range of complex well-posed problems in pure and applied mathematics, making productive use of knowledge and problem solving strategies."
Claim #3 - Communicating Reasoning	"Students can clearly and precisely construct viable arguments to support their own reasoning and to critique the reasoning of others."
Claim #4 - Modeling and Data Analysis	"Students can analyze complex, real-world scenarios and can construct and use mathematical models to interpret and solve problems."
Overall Claim for Grades 3-8	"Students can demonstrate progress toward college and career readiness in mathematics."
Overall Claim for Grade 11	"Students can demonstrate college and career readiness in mathematics."

Smarter Balanced
Assessment Consortium

Maintaining High Item Quality

Dates Subject to change

Sample Items for Smarter Balanced Assessments

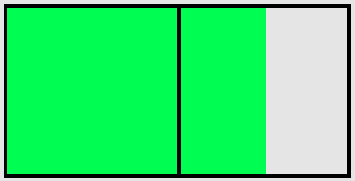
- How did the development of sample items inform our next steps as a consortium? What did we learn from the process?

http://sampleitems.smarterbalanced.org/itempreview/sbac/

From the Progressions

3.NF.1 Understand a fraction $1/b$ as the quantity formed by 1 part when a whole is partitioned into b equal parts; understand a fraction a/b as the quantity formed by a parts of size $1/b$.

The importance of specifying the whole



Without specifying the whole it is not reasonable to ask what fraction is represented by the shaded area. If the left square is the whole, the shaded area represents the fraction $\frac{3}{2}$; if the entire rectangle is the whole, the shaded area represents $\frac{3}{4}$.

Smarter Balanced Assessment Consortium

How do we “Shift” Item Writers Understanding of the Mathematics?

What fraction is represented by the red area?

Disclaimer: This is not a Smarter Balanced item.

Grow Capacity Through Many, Many Examples and Continuous Professional Learning

What fraction is represented by the red area?

Four students give responses. Explain what must be true for each student to be correct.

Student A: ¾

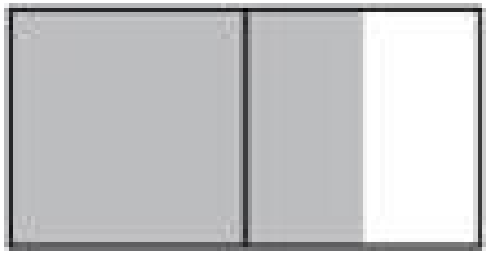
Student B: 3/2

Student C: 3/1

Student D: 30/1

Fractions Example

Look at the fraction model shown.

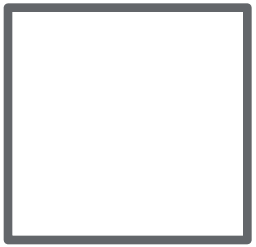


The shaded area represents 3/2. Drag the figures below to make a model that represents 3 x 3/2.

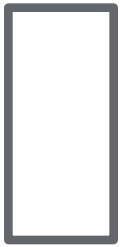
A

B

C



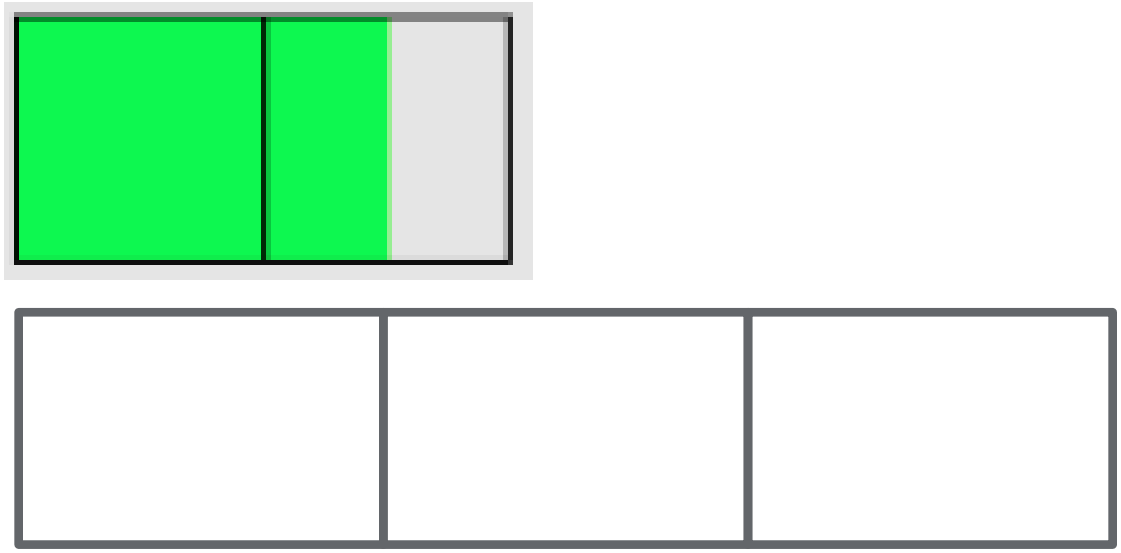
D



Smarter Balanced Assessment Consortium

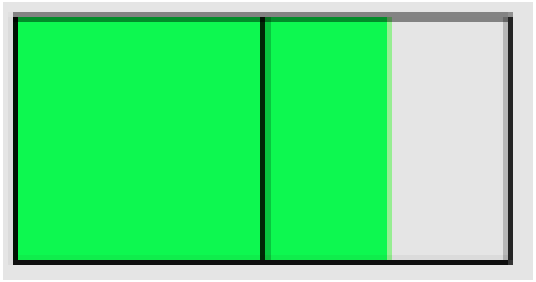
Student A drags three of shape B, which is equal in area to the shaded region. This student probably has good understanding of cluster 5.NF.B he knows that 3 x 3/2 is equal to 3 iterations of 3/2. Calculation of the product is not necessary because of the sophisticated understanding of multiplication.

Apply and extend previous understandings of multiplication and division to multiply and divide fractions.

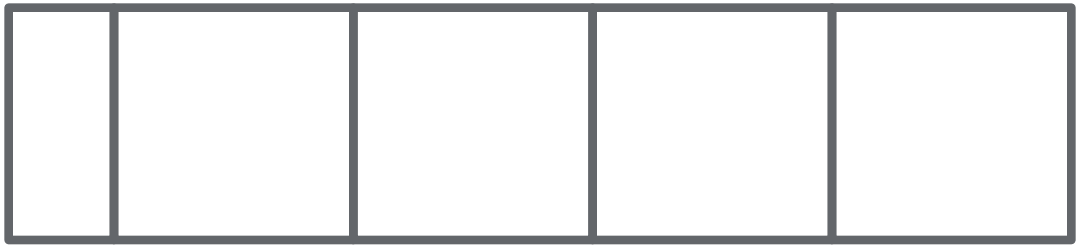


Student B reasons that 3 x 3/2 = 9/2 = 4 ½. She correctly reasons that since the shaded area is equal to 3/2, the square is equal to one whole, and drags 4 wholes plus half of one whole to represent the mixed number.

Apply and extend previous understandings of multiplication and division to multiply and divide fractions.

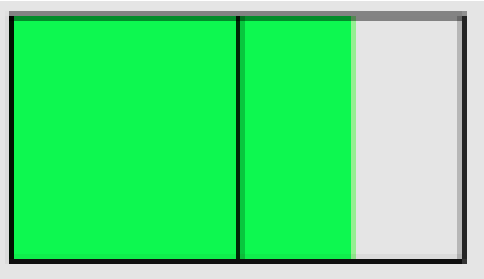


Note that unlike the previous chain of reasoning, this requires that the student determines how much of the shaded area is equal to 1.

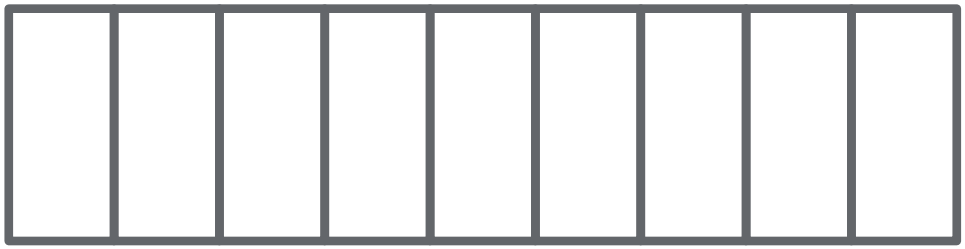


Student C multiplies 3 x 3/2 = 9/2. She reasons that since the shaded area is 3/2, this is equal to 3 pieces of size ½. Since 9/2 is 9 pieces of size ½, she drags nine of the smallest figure to create her model.

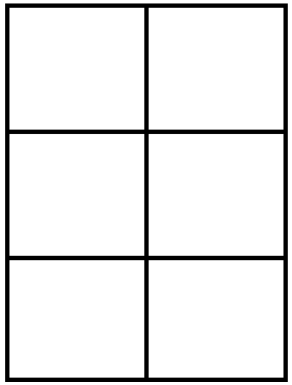
Apply and extend previous understandings of multiplication and division to multiply and divide fractions.



This chain of reasoning links nicely back to the initial development of 3/2 in 3.NF.1 "understand a fraction *a*/*b* as the quantity formed by *a* parts of size 1/*b*, illustrating the coherence in the standards across grades 3-5.

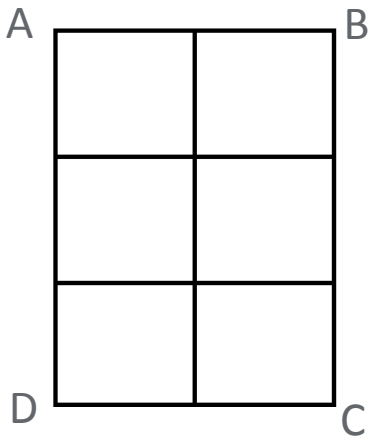


Give the area of the figure in square units.



Disclaimer: This is not a Smarter Balanced item.

The area of Rectangle ABCD is 24 square units. Draw a picture of 1 square unit.



Assessments Must Attend to the Coherence in the Standards

1.OA.7 Understand the meaning of the equal sign, and determine if equations involving addition and subtraction are true or false. *For example, which of the following equations are true and which are false? 6 = 6, 7 = 8 – 1, 5 + 2 = 2 + 5, 4 + 1 = 5 + 2.*

What it Looks Like in Grade 3

True or False:

3 x 8 = 20 + 4 T F

50 ÷ 10 = 5 x 1 T F

9 x 9 = 8 x 10 T F

What it Looks Like in Grade 5

True or False:

$$\frac{1}{2} \times \frac{1}{3} = \frac{3}{6} \times \frac{1}{3}$$

$$\frac{2}{2} \times \frac{1}{3} = \frac{3}{6} \times \frac{1}{3}$$

$$\frac{1}{2} \times \frac{1}{3} = \frac{\square}{6} \times \frac{1}{3}$$

What it Looks Like in Grade 8

Solve for x.

$3x + 17 = 3x + 12$

What it Looks Like in High School

$X^4 - 5x^3 + x^2 + 2x + 1 =$

Drag the correct expression to make a true equation.

$x^3 + (x + 1)^2 + X^4 - 6x^3$

$X^4 - 3x^3 + 2x^3 + x^2 + 2x + 1$

$X^4 - 5x^3 + x + x + 2x + 1$

…

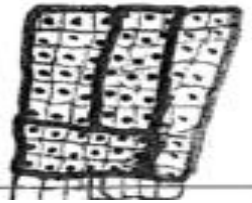
The Response Capture and Scoring Challenge

Dru and Teller had a tent that is 8 feet-by-10 feet.
Each adult has a sleeping bag that is 3-feet by 6-feet.

a. Dru said that four adults would fit in the tent. Each adult needs 18 square feet of floor space. 18 + 18 + 18 + 18 = 72. The tent is 80 square feet, so there is room to spare. Teller said that he tried and could not get four adults to fit in Tent C.

Who is right? Teller

b. Explain. Teller is right because you can fit three but there is more sq.ft. Left over enough to equal 18 just it wasn't a 3x6 it was a 9x2



Green = 9x2
Red = 3x6

Gavin, M. K., Casa, T. M., Chapin, S., Copley, J. V., & Sheffield, L. J. (2008). Project M2: Using Everyday Measures: Measuring with the Meerkats from Project M2: Mentoring Young Mathematicians series.

Bestsize Cans

The Fresha Drink Company is marketing a new soft drink.

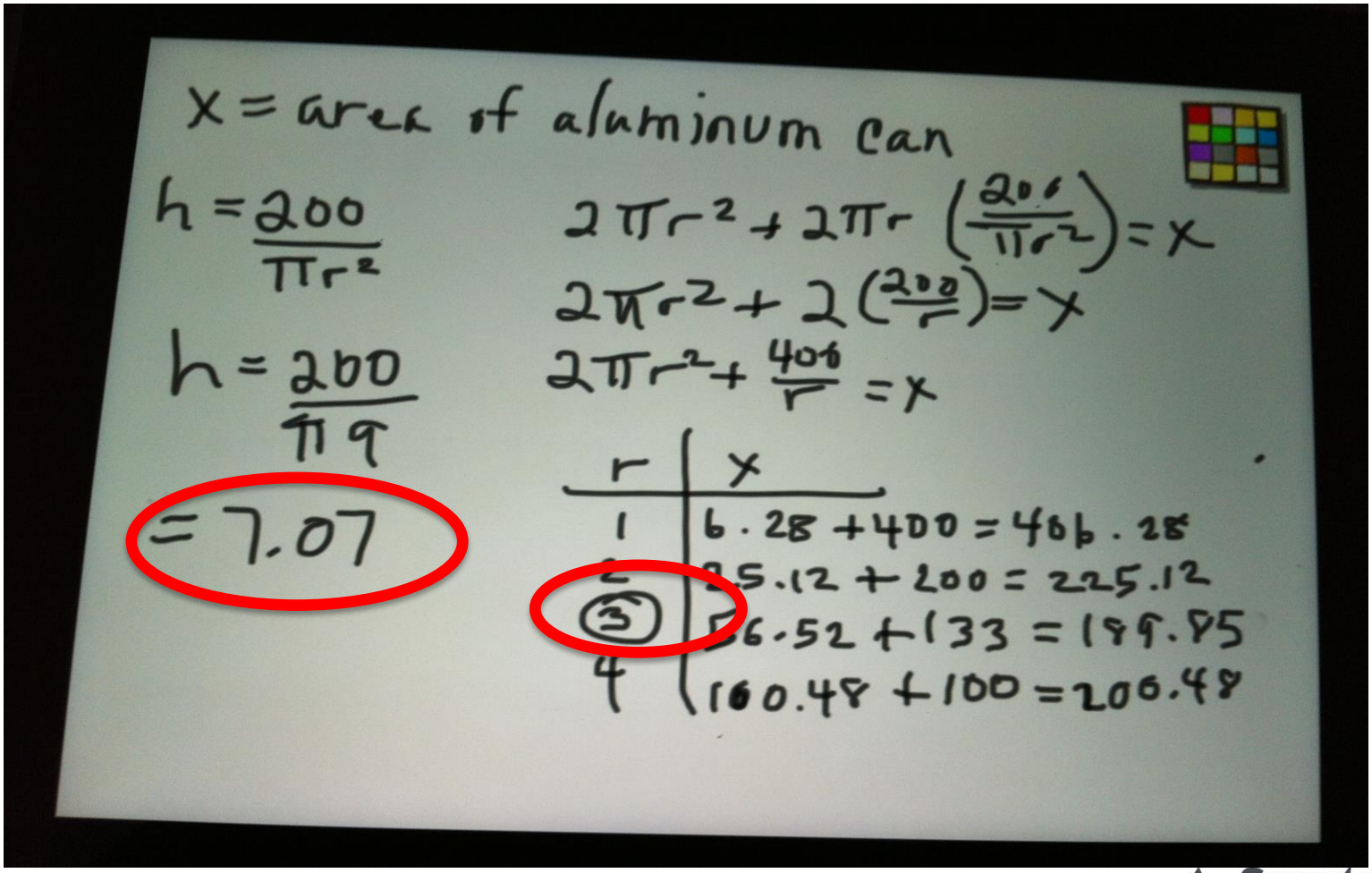
The drink will be sold in a can that holds 200 cm^3.

In order to keep costs low, the company wants to use the smallest amount of aluminum.

Find the radius and height of a cylindrical can which holds 200 cm3 and uses the smallest amount of aluminum.

Explain your reasons and show all your calculations

Response Capture is Not the Real Issue



The Ideal Sequence of Item Development, Response Capture, Scoring

Put the Math First

Create items aligned to the CCSS and Smarter Balanced Content Specifications regardless of real or perceived constraints.

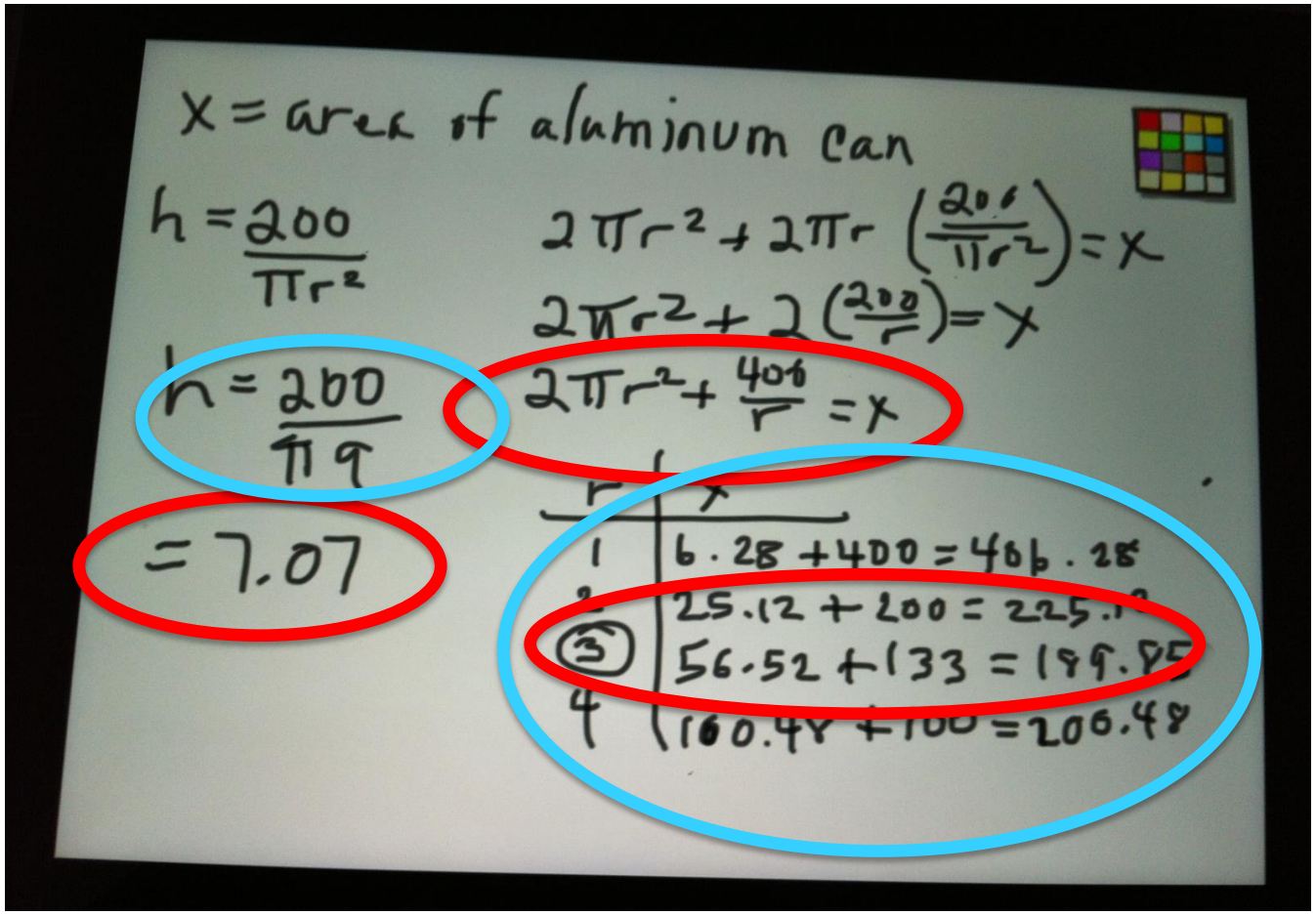
Natural Response Format

Determine the most effective way to capture the variety of expected student responses for a particular item or class of items using available technology.

Pushing the Field on AI Scoring

Score responses using detailed rubrics and maximize capabilities of AI through increasingly sophisticated data mining techniques.

Smarter Balanced Assessment Consortium


X = area of aluminum can
h = 200/πr²
2πr² + 2πr (200/πr²) = X
2πr² + 2(200/r) = X
h = 200/π9
2πr² + 400/r = X
= 7.07
r | X
1 | 6.28 + 400 = 406.28
3 | 56.52 + 133 = 189.52
4 | 160.48 + 100 = 206.48

Recent and Upcoming Developments

Smarter Balanced
Assessment Consortium

Recent and Upcoming Developments
(www.smarterbalanced.org)

- Early October
  - Release of 50 **sample Items and Tasks**, using prototype online delivery system

- Mid November
  - Adoption of Summative **Assessment Blueprint**
    - Allocation of points
    - Testing times
    - Breadth and depth of coverage of Common Core

Smarter Balanced Assessment Consortium

Recent and Upcoming Developments
(www.smarterbalanced.org)

- Late November
  - Release of draft **Achievement Level Descriptors** (ALDs) for review/comment
    - Open for comments through January 15
    - ALDs for all tested grades
    - Governing State adoption in March

- Early December
  - **Minimum technology specifications:** Hardware, operating systems, and bandwidth

Smarter Balanced Assessment Consortium

Recent and Upcoming Developments
(www.smarterbalanced.org)

- Mid January
  - Completion of **Validity Framework** for Smarter Balanced
    - Lays out research agenda to validate assessment claims and Theory of Action

- Mid February (Feb. 20)
  - Began **Pilot Testing** of first 5,000 items/tasks
    - "Scientific Sample" of about 10% of students
    - Practice tests at all grades conforming to blueprint released to public late May

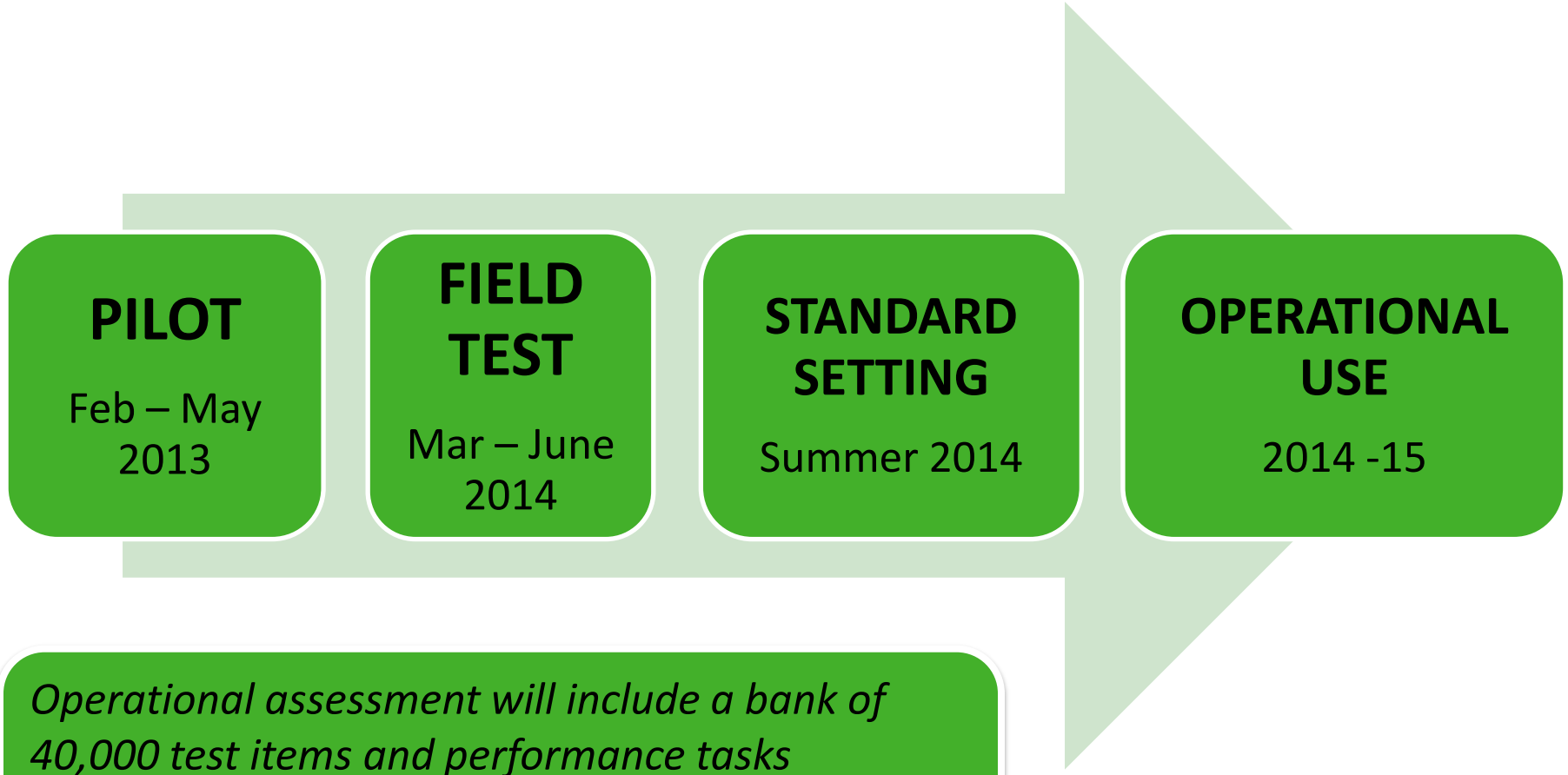
Smarter Balanced Assessment Consortium

Pilot Testing

- Began February 20; ends May 24
- Sample of >1M students (> twice as many as NAEP)
- Participation of more than 5,000 schools
- Purpose: Evaluate the efficacy of our first 5,000 items & tasks…
  - ✓ Do our Performance Tasks that involve real-world problem solving work well?
  - ✓ Can we computer-score open-ended questions?
  - ✓ Can students use the online tools?
  - ✓ Have we avoided bias in our items/tasks?
  - ✓ How well are we measuring grade-to-grade growth?

Smarter Balanced Assessment Consortium

Major Upcoming Milestones for the Summative Assessment



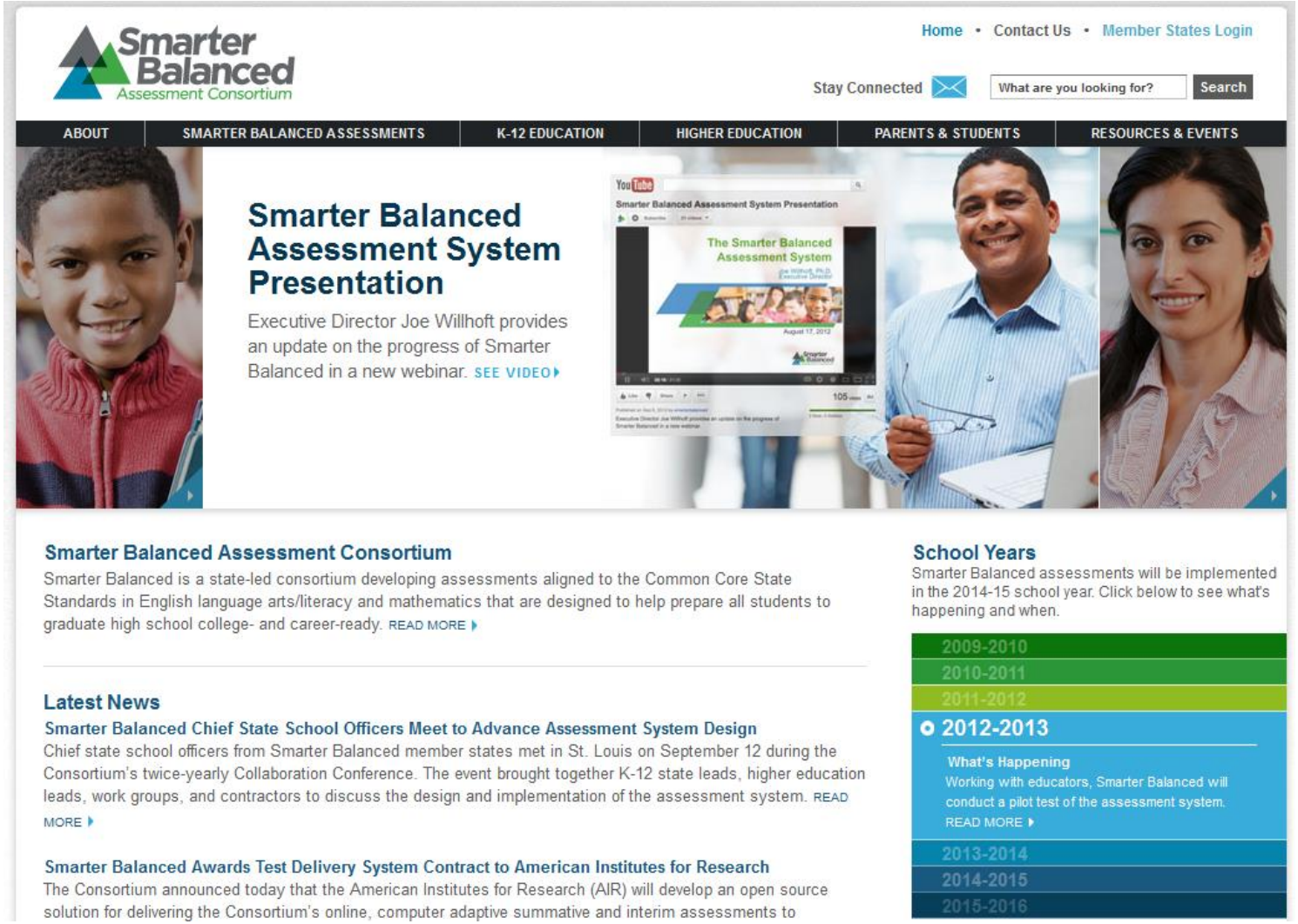
Operational assessment will include a bank of 40,000 test items and performance tasks administered to 9 million students across 21 states.

Smarter Balanced Assessment Consortium

Summary

- What you see in 2015 will be far better than what you currently see
  - proof by contradiction for those who initially thought the consortium couldn't pull it off
- Putting a pilot test out early has taught us a lot about process and how to improve our process
- We will continue to put the mathematics first and figure out the supporting components of delivery, response capture, and scoring based on what the math requires

Visit us at: SmarterBalanced.org



Smarter Balanced
Assessment Consortium