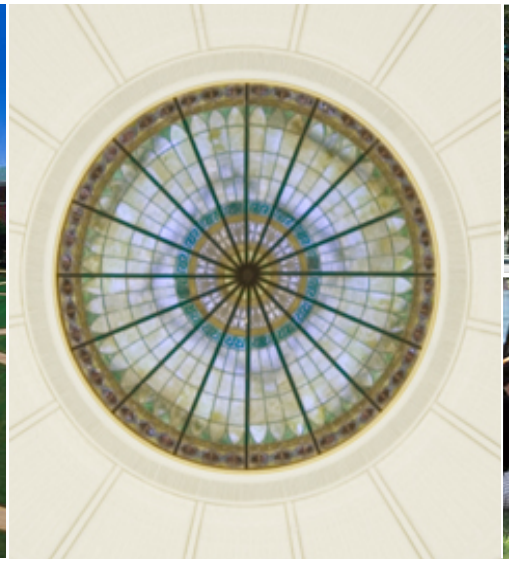
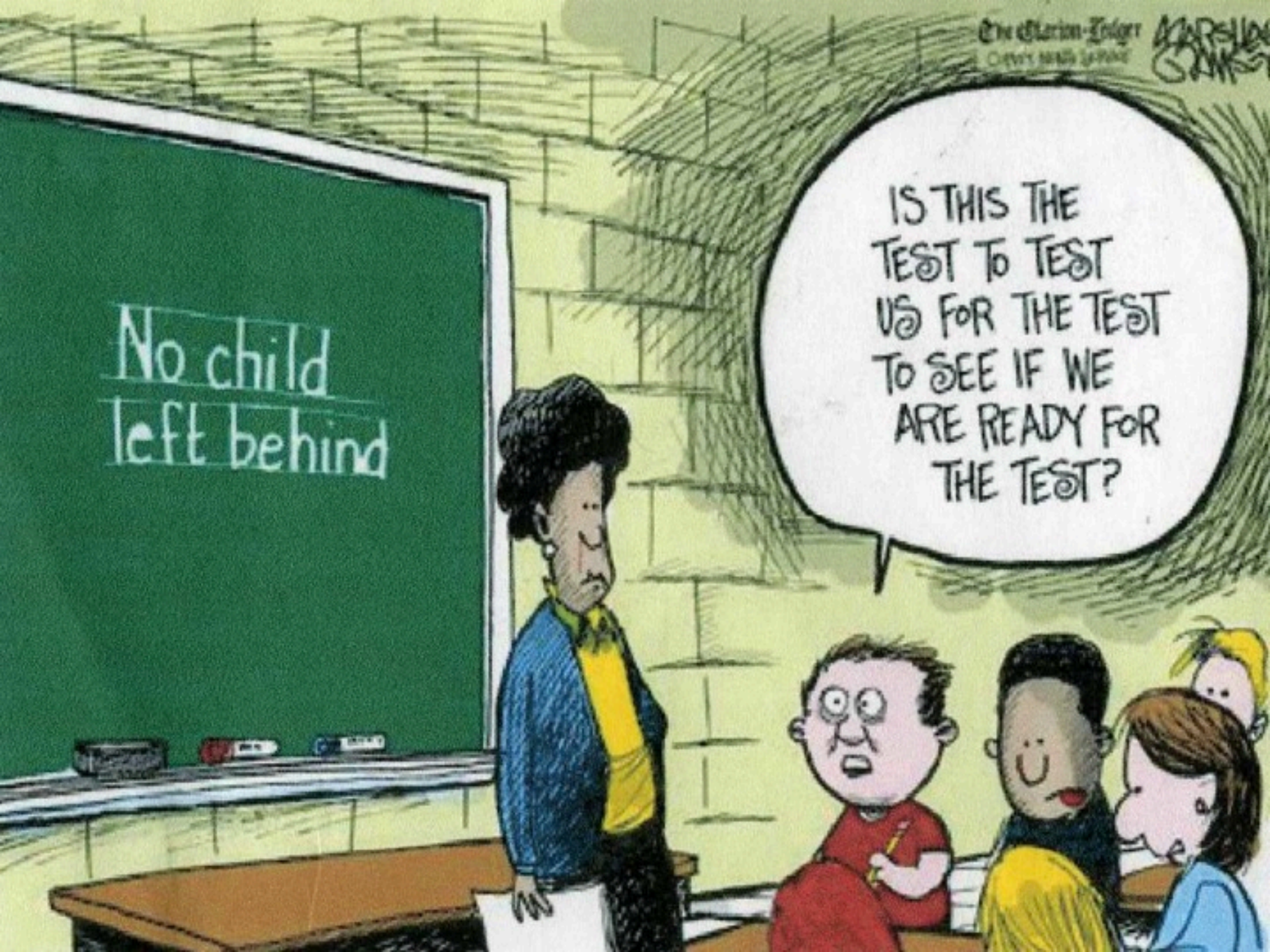



The Anatomy of High-Quality Multiple Choice Assessment Items

Research in Mathematics Education

Cassandra Hatfield


No child left behind
IS THIS THE TEST TO TEST US FOR THE TEST TO SEE IF WE ARE READY FOR THE TEST?

Focus on Assessment

Discussion Points	Outcome
• Knowledge Representations • Levels of Mathematics Proficiency • Appropriate Question Stems • Including Student Misconceptions in answer choices	• Understand the anatomy of a test item
• Importance of technically adequate assessments • Examining student performance: Moving beyond the overall score	• Reliable data decision making

Item Writing Template

Item Writing Template						
Grade Level:	TEKS:					
Knowledge Representation (Circle One):		**Foundational**		**Bridging**		**Target**
Cognitive Engagement (Circle One):	**Procedural**		**Conceptual**		**Strategic**	**Adaptive**
Relative Difficulty (Circle One):			**Easy**	**Medium**		**Difficulty**
Stem (Text and Graphic)	**Response Choices**					
		Content/Graphic			**Student Misconception Or Error**	
	Correct					
	Distractor 1					
	Distractor 2					
	Distractor 3					
Approval: _____ _____ _____ _____ _____ _____ **Writer Reviewer Writer Reviewer Writer Reviewer**	Reviewer Comments:					

Considerations for Test Items

- **Knowledge Representations**
  - Target knowledge and skills
  - Bridging knowledge and skills
  - Foundational knowledge and skills
- **Strands of Mathematical Proficiency**
  - Conceptual understanding
  - Procedural fluency
  - Strategic competence
  - Adaptive reasoning

Purpose of Different Types of Assessments

Universal Screening

- Who is struggling?
- What level of intensity of support does the student need?
- 3-Tier Response to Intervention Model

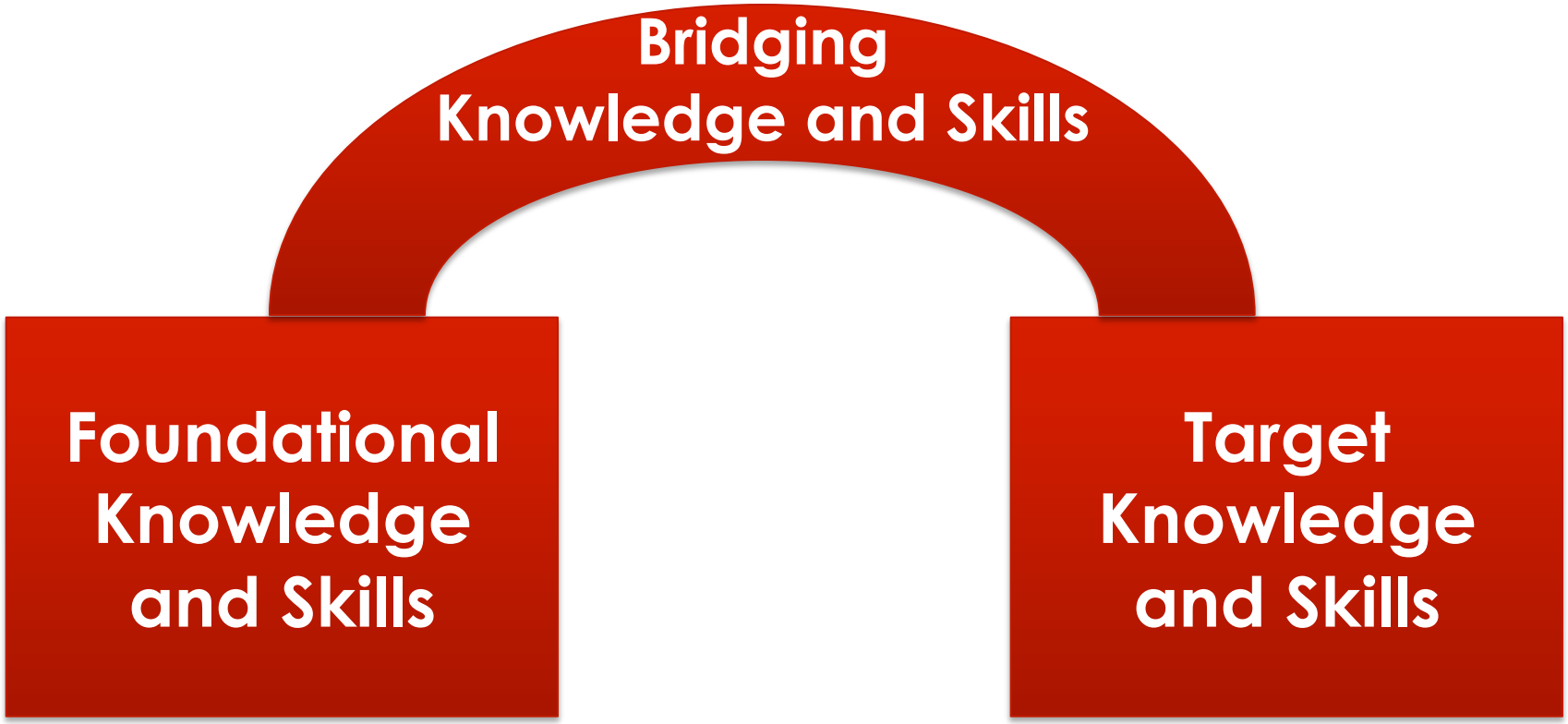
Diagnostic Assessments

- Why are they struggling?
- How deep are the gaps in the students understanding?
- Learning Progressions

Skills Inventory

- Why are they struggling?
- What grade level concepts do I need to intervene with to get them back on track?
- Individual Standards

Knowledge Representations



Knowledge Representations: Target Knowledge and Skills

- Grade level knowledge and skills
- Supports **future success in mathematics**
- Often are abstract representations of formal mathematical knowledge, but not always

Target Knowledge and Skills

Knowledge Representations: Foundational Knowledge and Skills

- Knowledge and skills that support the target content and are accumulated from previous learning
- **Prerequisite knowledge and skills** needed to be successful at the target knowledge and skills

Foundational Knowledge and Skills

Knowledge Representations: Bridging Knowledge and Skills

- Knowledge and skills needed to **connect or support students' learning** from the foundational knowledge to the target knowledge and skills
  - Often represents an integration of knowledge and skills (may be conceptual or model-based)
- The knowledge and skills that students learn from the teacher or instructional materials

Sample **Foundational** Item
Grade 3

Gracie has 12 books. She has an empty bookshelf with 3 shelves.

She puts the same number of books on each shelf, how many books will be on each shelf?

A. 3 books

B. 4 books

C. 6 books

D. 12 books

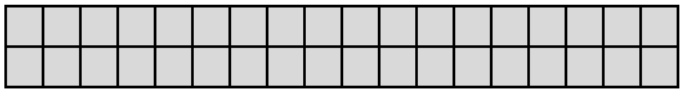
Correct answer: **B**

Sample **Bridging** Item
Grade 3

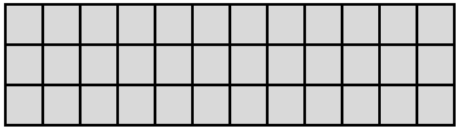
Which model could be used to represent

$36 \div 9 = \square$?

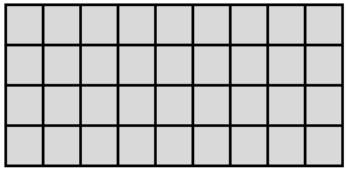
A.



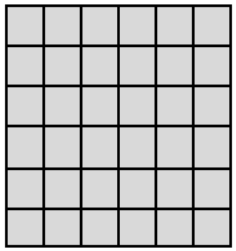
B.



C.



D.



Correct answer: **C**

Connecting Knowledge Representations with the TX-RCFP

Curriculum Focal Points are:

- Framework of critical areas of mathematics instruction
- A mathematical theme, not a single TEKS statement

Before writing assessments, it is important to:

- Analyze the focal point description to identify key target skills
- Analyze related TEKS statements
- Synthesize the overlaying skills

TEXAS RESPONSE TO CURRICULUM FOCAL POINTS FOR GRADE 3 MATHEMATICS REVISED 2013

Understanding and applying place value and properties of operations to solve problems involving addition and subtraction of whole numbers within 1,000

Students extend their understanding of the base-10 system to numbers up to 100,000 and represent addition and subtraction of numbers within 1,000 using pictorial models, number lines, and equations. They use efficient, accurate, and generalizable methods based on place value, properties of operations, and the relationship between addition and subtraction to solve problems involving addition and subtraction of whole numbers within 1,000.

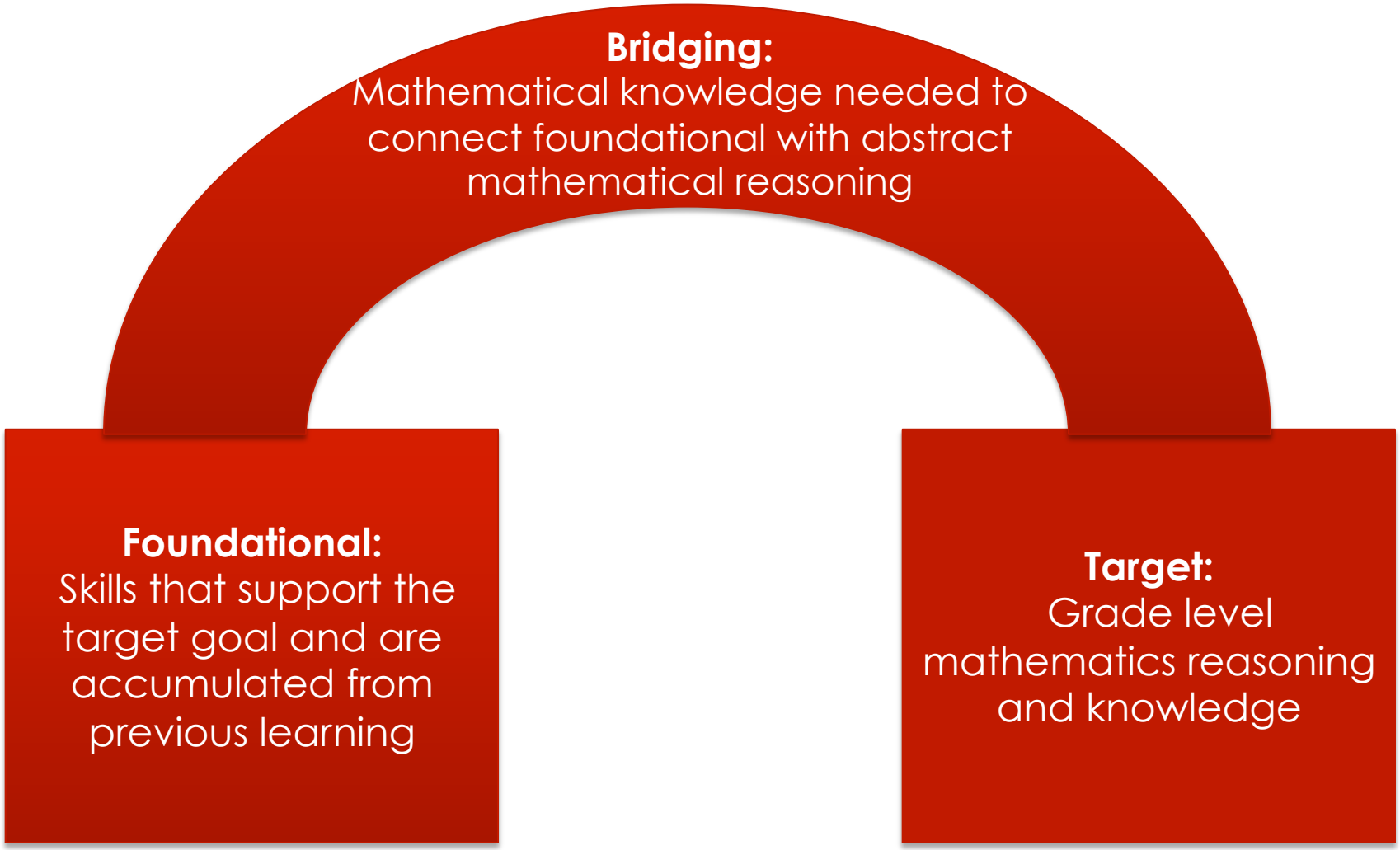
Related Grade 3 TEKS:

3(1)(A)-(G)	The student uses mathematical processes to acquire and demonstrate mathematical understanding.
3(2)(A)	The student is expected to compose and decompose numbers up to 100,000 as a sum of so many ten thousands, so many thousands, so many hundreds, so many tens, and so many ones using objects, pictorial models, and numbers, including expanded notation as appropriate.
3(2)(B)	The student is expected to describe the mathematical relationships found in the base-10 place value system through the hundred thousands place.
3(2)(C)	The student is expected to represent a number on a number line as being between two consecutive multiples of 10; 100; 1,000; or 10,000 and use words to describe relative size of numbers in order to round whole numbers.
3(2)(D)	The student is expected to compare and order whole numbers up to 100,000 and represent comparisons using the symbols >, <, or =.
3(4)(A)	The student is expected to solve with fluency one-step and two-step [multi-step] problems involving addition and subtraction within 1,000 using strategies based on place value, properties of operations, and the relationship between addition and subtraction.
3(4)(B)	The student is expected to round to the nearest 10 or 100 or use compatible numbers to estimate solutions to addition and subtraction problems.
3(4)(C)	The student is expected to determine the value of a collection of coins and bills.
3(5)(A)	The student is expected to represent one- and two-step problems involving addition and subtraction of whole numbers to 1,000 using pictorial models, number lines, and equations.
3(7)(C)	The student is expected to determine the solutions to problems involving addition and subtraction of time intervals in minutes using pictorial models or tools such as a 15-minute event plus a 30-minute event equals 45 minutes.
3(8)(A)	The student is expected to summarize a data set with multiple categories using a frequency table, dot plot, pictograph, or bar graph with scaled intervals.
3(8)(B)	The student is expected to solve one- and two-step problems using categorical data represented with a frequency table, dot plot, pictograph, or bar graph with scaled intervals.

Process for Articulating the Content of an Item through knowledge representations

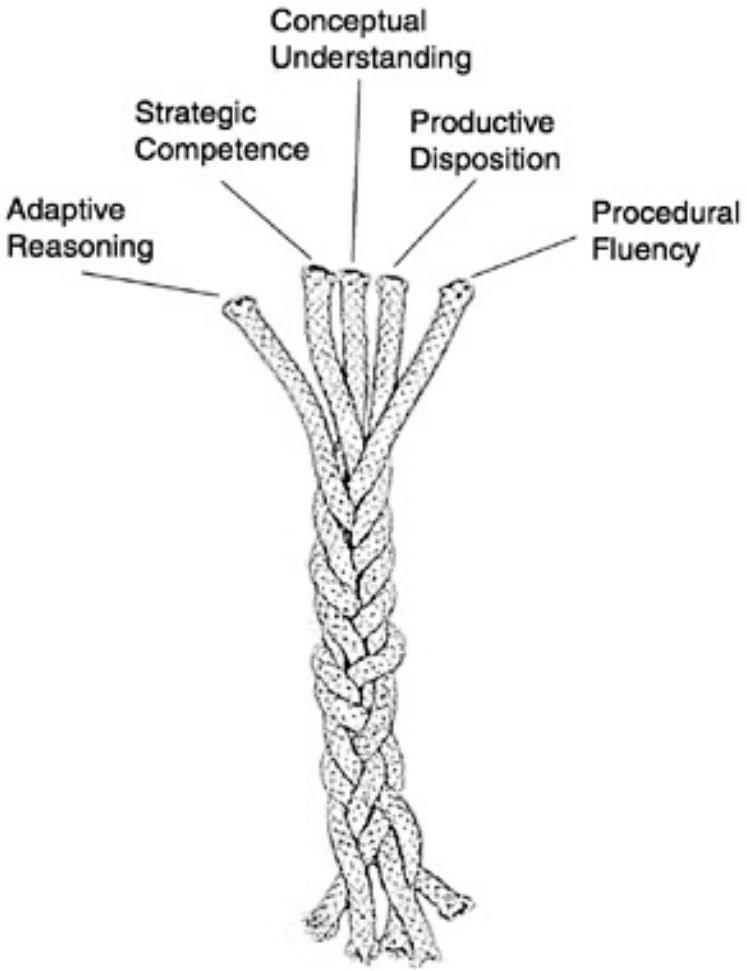
- Become familiar with the TEKS standard (content standard) for which you are writing a test item
- Articulate the **TARGET SKILLS**
- Articulate the **FOUNDATIONAL SKILLS**
- Articulate the **BRIDGING SKILLS**

Knowledge Representations



Strands of Mathematical Proficiency

- Intertwined Strands of Proficiency



National Research Council (2001)

Conceptual Understanding

- Demonstrate an integrated and functional grasp of mathematical ideas
- Understand specific task as it relates to a whole concept
- Find relationships between pieces of information
- Make connections to similar representations
- Use models and multiple representations (e.g. pictures, numbers, real-life situations, words)

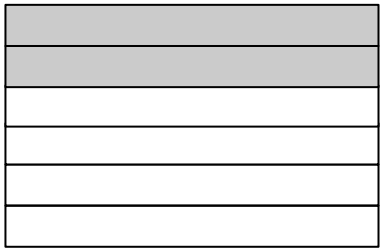
Conceptual Understanding

TEKS 6.4E

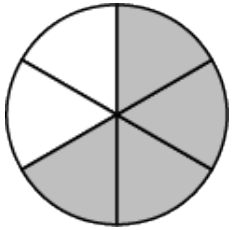
The student applies mathematical process standards to develop an understanding of proportional relationships in problem situations. The student is expected to:

(E) **Represent ratios and percents with concrete models, fractions, and decimals.**

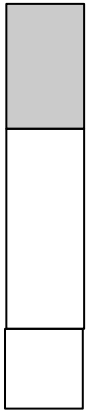
Which model has an equivalent ratio of shaded parts to total parts as this rectangle?



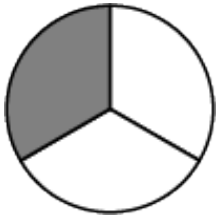
A



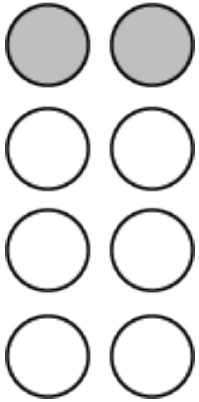
B



C



D



Correct answer: **C**

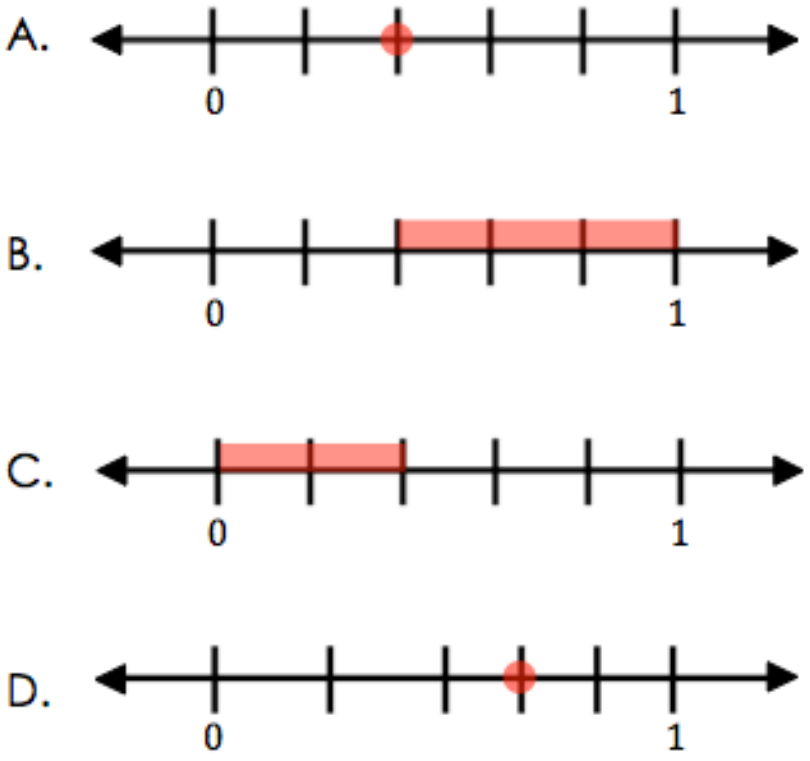
Conceptual Understanding

TEKS 3.3D

The student applies mathematical process standards to represent and explain fractional units. The student is expected to:

(D) **Compose and decompose a fraction a/b with a numerator greater than zero and less than or equal to b as a sum of parts $1/b$.**

Which number line shows $\frac{3}{5}$?



Correct answer: **B**

Procedural Fluency

- Use formal language or symbolic representations
- Carry out accurate computations
- Follow multiple steps sequentially
- Make proper use of algorithm and properties

Procedural Fluency

TEKS 7.11A

The student applies mathematical process standards to solve one-variable equations and inequalities. The student is expected to:

(A) **Model and solve one-variable, two-step equations and inequalities.**

Solve for w:

$$\frac{w + 12}{3} = 20$$

A. 72

B. 58

C. 48

D. 16

Correct answer: **C**

Procedural Fluency

TEKS 4.3E

The student applies mathematical process standards to represent and generate fractions to solve problems. The student is expected to:

(E) **Represent and solve addition and subtraction of fractions with equal denominators using objects and pictorial models that build to the number line and properties of operations.**

Mario cuts a pie into 6 equal slices. He eats 1 slice for lunch and 1 slice for dinner. Which equation shows the fraction of a whole pizza that Mario eats?

A $\frac{1}{6} + \frac{1}{6} = \frac{1}{12}$

B $\frac{1}{6} + \frac{1}{6} = \frac{2}{12}$

C $\frac{1}{6} + \frac{1}{6} = \frac{2}{6}$

D $\frac{1}{3} + \frac{1}{3} = \frac{2}{6}$

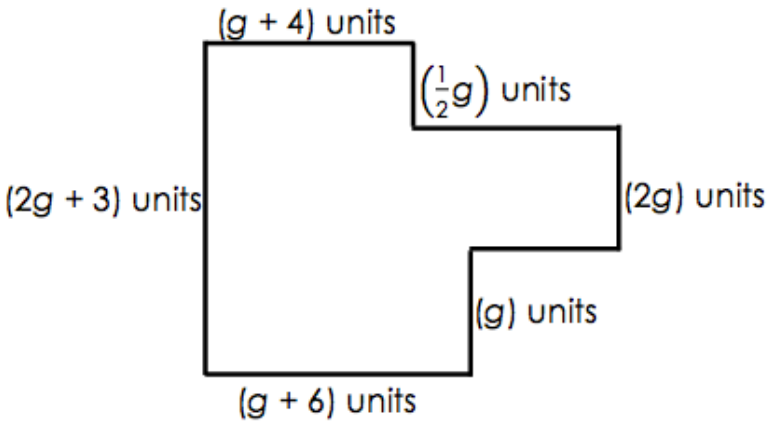
Correct answer: **C**

Strategic Competence

- Ability to formulate a problem in mathematical terms
- Represent problem solving strategically (verbally, symbolically, graphically, or numerically)
- Identify and use strategy necessary to solve problems effectively (e.g. use the distributive property to solve)

Strategic Competence

The perimeter of the figure below is 43 units.



Which equation can be used to solve for the variable, g?

A. $7.5g + 13 = 43$

B. $7g + 13 = 43$

C. $5.5g + 13 = 43$

D. $5g + 13 = 43$

Correct answer: **A**

Strategic Competence

TEKS 3.4A

The student applies mathematical process standards to develop and use strategies and methods for whole number computations in order to solve problems with efficiency and accuracy. The student is expected to:

(A) **Solve with fluency one-step and two-step problems involving addition and subtraction within 1,000 using strategies based on place value, properties of operations, and the relationship between addition and subtraction.**

Samantha takes $140 to the toy store. She buys a doll house and has $58 left over. Which number sentence can be used to find the cost of the doll house?

A $\blacksquare - 140 = 58$

B $140 + 58 = \blacksquare$

C $140 + \blacksquare = 58$

D $140 - \blacksquare = 58$

Correct answer: **D**

Adaptive Reasoning

- Think logically about a problem, which requires reflecting on various approaches to solve a problem and deductively selecting an approach
- Rationalize and justify strategies
- Appropriately explain a procedure or concept

Adaptive Reasoning

TEKS 3.3H

The student applies mathematical process standards to represent and explain fractional units. The student is expected to:

(H) **Compare two fractions having the same numerator or denominator in problems by reasoning about their sizes and justifying the conclusion using symbols, words, objects, and pictorial models.**

Why is $\frac{2}{3} > \frac{2}{7}$?

A. 2 is equal to 2 and thirds are larger than sevenths.

B. 2 is equal to 2 and thirds are smaller than sevenths.

C. 2 is equal to 2 and 3 is smaller than 7.

D. 2 is equal to 2 and 3 is greater than 7.

Correct answer: **A**

Adaptive Reasoning

TEKS 6.2B

The student applies mathematical process standards to represent and use rational numbers in a variety of forms. The student is expected to:

(B) **Identify a number, its opposite, and its absolute value.**

Adam says that the absolute value of -7 Is 7. Why is he correct?

A. 7 is the distance from -7 to 0 on the number line.

B. 7 is a prime number with only two factors.

C. 7 is an odd number, so it is not divisible by 2.

D. 7 is the square root of 49, which is a perfect square.

Correct answer: **A**

Level of Difficulty

- Easy
  - Basic Knowledge
  - Skills that are familiar to students
  - Sometimes conceptually based
- Medium
- Difficult
  - Skills that are peripheral to curriculum
  - Not all students will have mastered these

(Cheng, 2005)

Item Match

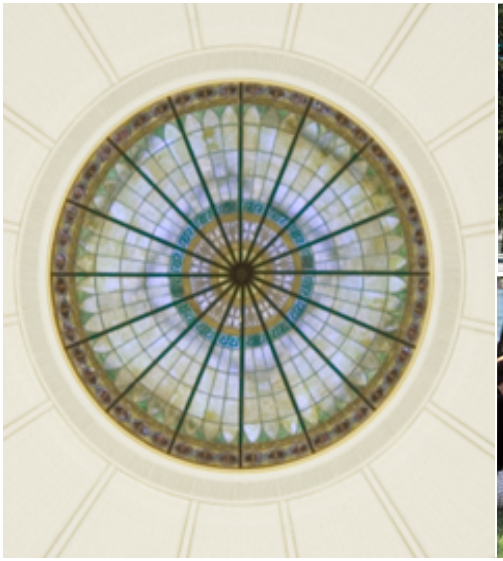
- In your packet of materials is a 4 x 3 matrix with the 4 strands of mathematical proficiency along the top and 3 levels of difficulty along the left side.
- Take a blue set of cards. They should have a number in the corner, ranging from 1 to 12.
- As you read each item, determine the strand of mathematical proficiency and relative level of difficulty.
- Write the number associated with the item in the appropriate cell.

Item Match – Answer Key

	Procedural	Conceptual	Strategic	Adaptive
Easy	5	3	1	12
Medium	10	6	4	7
Difficult	8	11	9	2

Consider your purpose for giving the assessment when using these considerations

- **Knowledge Representations**
  - Target knowledge and skills
  - Bridging knowledge and skills
  - Foundational knowledge and skills
- **Strands of Mathematical Proficiency**
  - Conceptual understanding
  - Procedural fluency
  - Strategic competence
  - Adaptive reasoning

Assessment Item Development

General Item-Writing (Procedures)

- Avoid the complex multiple-choice format. (i.e., A and D, B and C).
- Use plain language. Avoid conditional phrases (if…, then…).
- Keep the language of the stem and response options at the appropriate grade level.
- Minimize examinee reading time in phrasing each item.

(Haladyna, 2004)

Stem Development

- State the stem in question form. Minimize use of completion form. When using the completion format, do not leave a blank for completion in the beginning or middle of the stem.
- Include only the material needed to make the problem clear. Avoid extraneous information.
- Word the stem positively; avoid negative phrasing. If an item must be stated negatively, underline or capitalize the negative word.
- Keep all essential information in the stem. Items that require students to read and evaluate each response option prior to selecting an answer increase the cognitive load required.

(Haladyna, 2004)

General Item-Writing (Content)

- Base each item on important content to learn; avoid trivial content.
- Keep the content of each item independent from content of other items on the test.
- Avoid cuing one item with another; keep items independent of one another.
- Avoid items based on opinions.
- Develop items that measure higher-level thinking.
- Avoid potentially insensitive content or language.

(Haladyna, 2004)

Response Development

- Make all distractors plausible.
  - If you're interested in obtaining more information about students' understanding, create distractors that represent common misconceptions may have about the content being assessed
- Keep all options in an item homogenous in content and grammatical structure.
- Keep the length of options brief and fairly consistent.
- Phrase options positively, not negatively.

(Haladyna, 2004)

Uh Oh ☹
Examples of Poorly Written Math Items

Dan enjoys going to the park. The park has lots of trees and a pond. Dan goes to the park at 10:15 a.m., plays on the swings, slides down the slide, and runs around the pond. He is hungry when he finally makes it home 1 hour 55 minutes later. What time did he arrive at home?

A. 12:00 p.m.

B. 12:10 p.m.

C. 8:20 a.m.

D. 6:00 p.m.

Uh Oh ☹
Examples of Poorly Written Math Items

Sue has a box of 2 stars and 2 circles. She wants to make her box proportional to Michelle's box of 6 stars and 3 circles. How many stars does she need to add to her box to make it proportional to Michelle's box?

A. 6
B. 0
C. 2
D. 20


Information
Action Plan

Purpose of Different Types of Assessments

Universal Screening

- Who is struggling?
- What level of intensity of support does the student need?
- 3-Tier Response to Intervention Model

Diagnostic Assessments

- Why are they struggling?
- How deep are the gaps in the students understanding?
- Learning Progressions

Skills Inventory

- Why are they struggling?
- What grade level concepts do I need to intervene with to get them back on track?
- Individual Standards

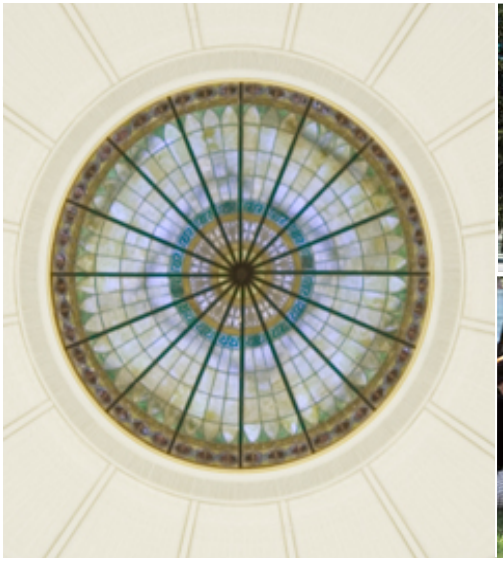
Decision Making Checklist for Assessment Development

Teachers should agree to analyze the assessments around the same set of criteria. The decisions should be directed toward:

- ✓ Validation about the appropriateness of the assessment
- ✓ Ensuring the assessment is congruent to the stated mastery objective and/or state or district standards
- ✓ Consistency of opinion about the assessment and evaluation of the work
- ✓ Adjustments in teacher directions and support for all students

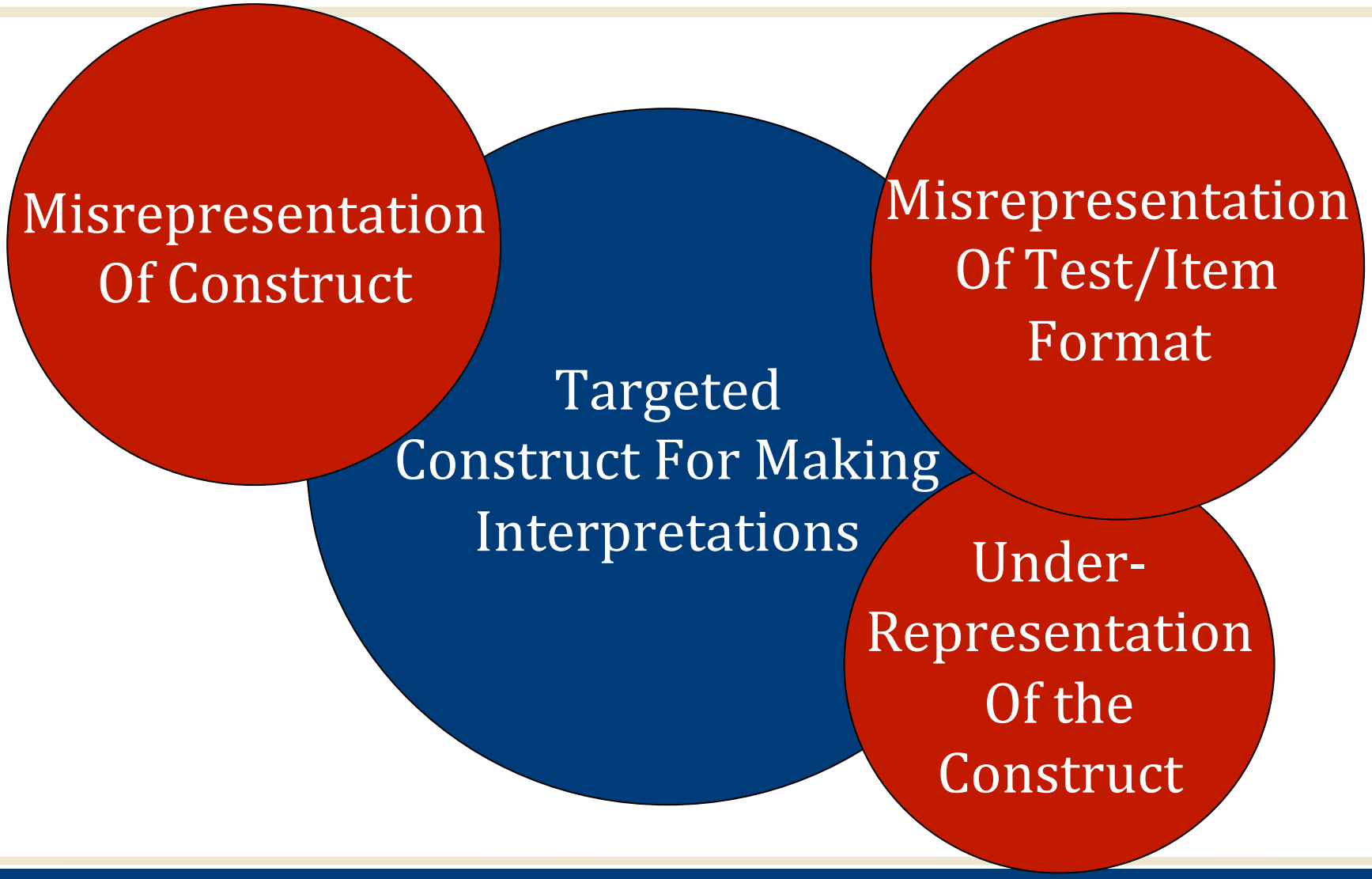
(Rutherford, 2008)

SMU



Data-Driven Decision Making

Barriers to Decision Making



Using Data to Inform Instruction: Overall Student Performance

- One of the most common ways to examine student performance data is by making normative comparisons of their **overall** performance:

Grade	Teacher	Student	Number Correct	Percent Correct
5	Richardson	Swan, B.	20/40	50%
5	Richardson	Black, J.	22/40	55%
5	Richardson	Cullen, E.	21/40	52.5%
5	Richardson	Newton, M.	20/40	50%
5	Hatfield	Everdeen, K.	37/40	92.5%
5	Hatfield	Hawthorne, G.	36/40	90%
5	Hatfield	Meelark, P.	37/40	92.5%
5	Hatfield	Abernathy, H.	35/40	87.5%

Using Data to Inform Instruction: Digging a Little Bit Deeper

- Although students' overall performance may be similar, this does not necessarily mean that they have similar levels of mastery on the assessed content:

Student	Number of Items Correct				Total Correct
	Number Line Structure	Magnitude as Distance	Part-to-Whole Relationships	Unit Fractions	
Swan, B.	7/10	6/10	5/10	2/10	20/40
Black, J.	5/10	7/10	4/10	6/10	22/40
Cullen, E.	8/10	5/10	3/10	5/10	21/40
Hale, J.	5/10	7/10	3/10	5/10	20/40

Using Data to Inform Instruction: Digging Even Deeper

- Similarly, even though students may have responded correctly to the same number of items within a given sub-level, again this does not necessarily mean they have mastered the same skills or have the same level of proficiency with the targeted skill

Student	Part-to-Whole Relationships (Items 1-10)										Number Correct
	1	2	3	4	5	6	7	8	9	10	
Swan, B.	1	0	1	1	0	1	0	0	0	1	5/10
Black, J.	1	1	0	0	1	1	0	0	0	0	4/10
Cullen, E.	1	1	0	0	0	1	0	0	0	0	3/10
Hale, J.	1	0	1	0	0	1	0	0	0	0	3/10

0 = Incorrect
1 = Correct

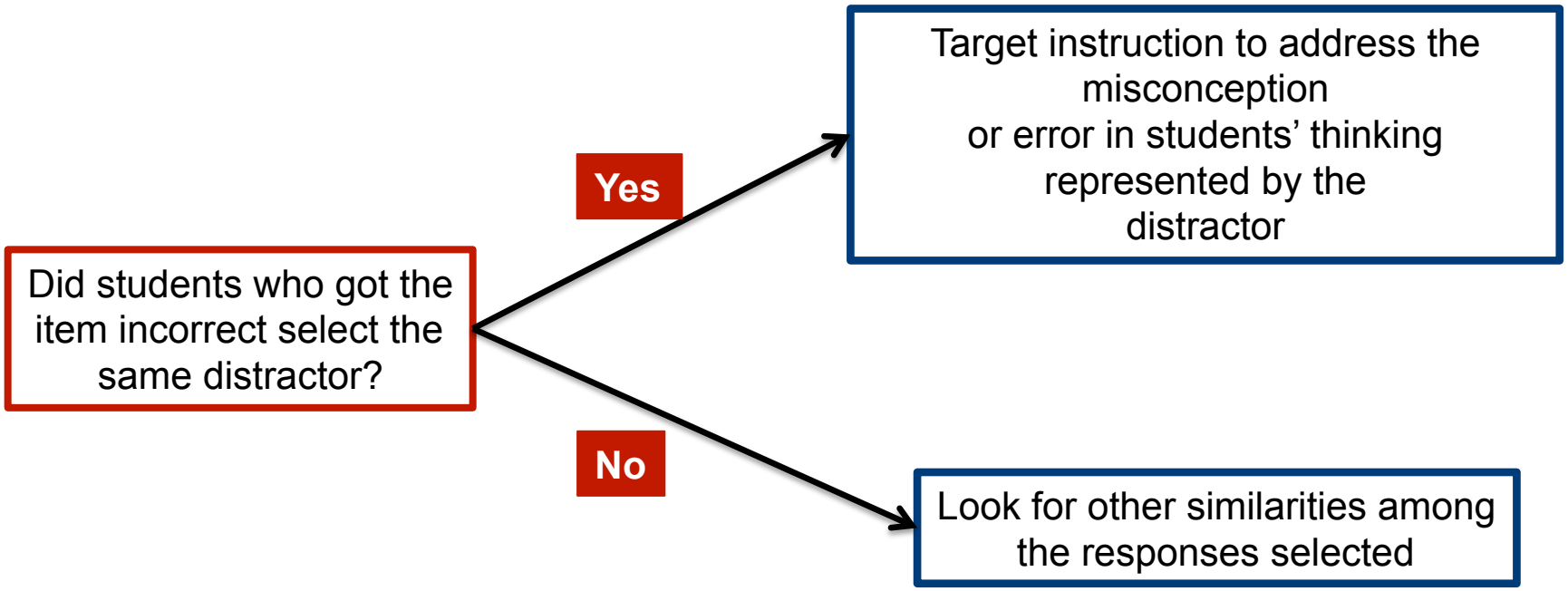
Using Data to Inform Instruction: How Far Down the Rabbit Hole Should I Go?

Clearly, the level at which you can examine student performance can become increasingly fine-grained. For example, once you start looking at item-level performance you can examine any of the following attributes:

- Whether students selected one distractor more frequently than the others (e.g., Did all students who got the item incorrect select the first distractor?)
- Whether students responded correctly to items targeting a specific level of proficiency (e.g., Did students get all of the items targeting procedural fluency correct?)
- Whether students consistently selected the incorrect response for items targeting a specific level of proficiency (e.g., Did students consistently get items targeting strategic competence or adaptive reasoning incorrect?)

Using Data to Inform Instruction: How Far Down the Rabbit Hole Should I Go?

Responses to these questions can then be used to further guide instructional planning. For example:



Using Data to Inform Instruction: How Far Down the Rabbit Hole Should I Go?

Did students respond correctly to all items targeting a certain level of cognitive complexity?

Yes → Consider varying the examples used during instruction and on assessments to include items that require more complex cognitive processing

No → Look for other similarities among the responses selected

Using Data to Inform Instruction: How Far Down the Rabbit Hole Should I Go?

Now it's time to consider the **BIG** question….

Do I need to look at the data this closely for every student in my classroom?

Our response….

Probably not. For students who are on track, monitoring their progress using the overall score is probably enough. For students who are struggling, however, digging a bit deeper into the data to try and figure out ***why*** they are struggling may be very useful.

Understanding ***why*** students are struggling can be used to help us target our instructional efforts to meet students' needs.

Some Final Thoughts & Take-Aways

- **Assessment and instruction should be considered together** – performance on assessments can inform instruction and assessments can be specifically designed to provide students with opportunities to demonstrate what they've learned during instruction.
- When considering the level of knowledge represented while designing a test item, start with the outcome (target knowledge) in mind. Then consider what **foundational knowledge** the student needs and the **bridging knowledge** that will help the student acquire the **target knowledge** or skill.

Some Final Thoughts & Take-Aways

- When designing tests or assessments for use in your classroom, be sure to include items that target **multiple levels of cognitive engagement**.
- Revisit the **guidelines for item development** as often as needed to ensure that the items you write provide students with the best opportunity possible to demonstrate their knowledge and understanding of the content.
- Don't hesitate to **look beyond the overall test score** to students' performance on groups of items or individual items to try and understand what instructional supports you can provide to support students' learning

References

- Leong, S.C. (2006). *On varying the difficulty of test items.* Paper presented at the 32nd Annual Conference of the International Association for Educational Assessment, Singapore.
- Haladyna, T. M. (2004). *Developing and validating multiple-choice items*. New York, NY: Routledge.
- National Research Council. (2001). *Adding it up: Helping children learn mathematics .* Washington, DC: National Academies Press.
- Rutherford, P. (2008) *Instruction for all students* (2nd ed.). Alexandria, VA: Just Ask Publications & Professional Development

Special thanks to…

- Deni Basaraba Ph.D
- Pooja Shivraj Ph.D.

Contact Information

www.smu.edu/RME

@RME_SMU

SMU Research in Mathematics Education- RME

Cassandra Hatfield
chatfield@smu.edu