

Considerations in Designing a
“2⁰% Assessment” (AA-MAS):
A beginning framework and examples
of conceptual possibilities

Brian Gong

Center for Assessment

Presentation at the Special Education Partnership
Conference on Alternate Assessments Based on
Modified Academic Achievement Standards

Washington, DC July 26, 2007

Regulation “Givens”

■ Purpose

- to ensure that all students with disabilities are counted in the accountability system, and are appropriately assessed.

■ Who

- Assessments for students whose disabilities preclude them from achieving grade level proficiency in the same timeframe as other students.
- Accountability – Limit on the use of scores for AYP to 2% of the total student population.



Regulation “Givens” - 2

- Intended Outcome
 - Since all students with disabilities are to be receiving instruction in the grade-level curriculum, these tests will not only ensure their inclusion in accountability systems, but also inform their instruction.
- What it must do
- What it should not do
 - **not OOLT (Out of (grade) level test)**



Instructional Vision

- Interaction with Instruction, How Administered
 - can be transitional/allows for valid assessment while accelerated growth makes up gap
 - not permanent status
 - annual decision
 - by content area
 - designed to convey BOTH concept of 'continuum' and notion of limited participation in alternates.
 - Linking Instruction and Assessment to State Standards
 - access to general curriculum (meaningful participation/opportunity to learn
 - standards-based IEPs; standards-based PLEPs, annual goals



My Purpose

- Provide *framework* and *examples* that states might find useful as they consider whether and how to design a “2% Assessment” (AA-MAS)
- Note: Work in progress
- Note: **Not** endorsed by USED or others
- May be useful in identifying issues for policy clarification of the regulation/state policy



Framework for Designing AA-MAS

- Within validity framework (see emerging Workbook for Technical Manual for AA-MAS)

- Define:

Purpose	Who Students Are
Construct/Domain	Proficiency
Reports (inferences, uses)	Assessment Specifications

- Consider technical and other aspects (practical, political)



Defining Purpose

- Assess students' knowledge and skills in relation to regular proficiency definition more accurately.
- Assess students' knowledge and skills in relation to a new proficiency definition.
- Provide incentives for educational system to provide more equitable and effective learning opportunities for these students, intended to result in their being proficient on the regular assessment to regular achievement standards.
- Provide incentive for educational system to provide social integration for students with disabilities, even though learning outcomes are not equivalent to regular students.
- Assess students' knowledge and skills more specifically (diagnose strengths and weaknesses)
- Provide a means for reducing school identification under AYP by declaring more students with disabilities "Proficient"



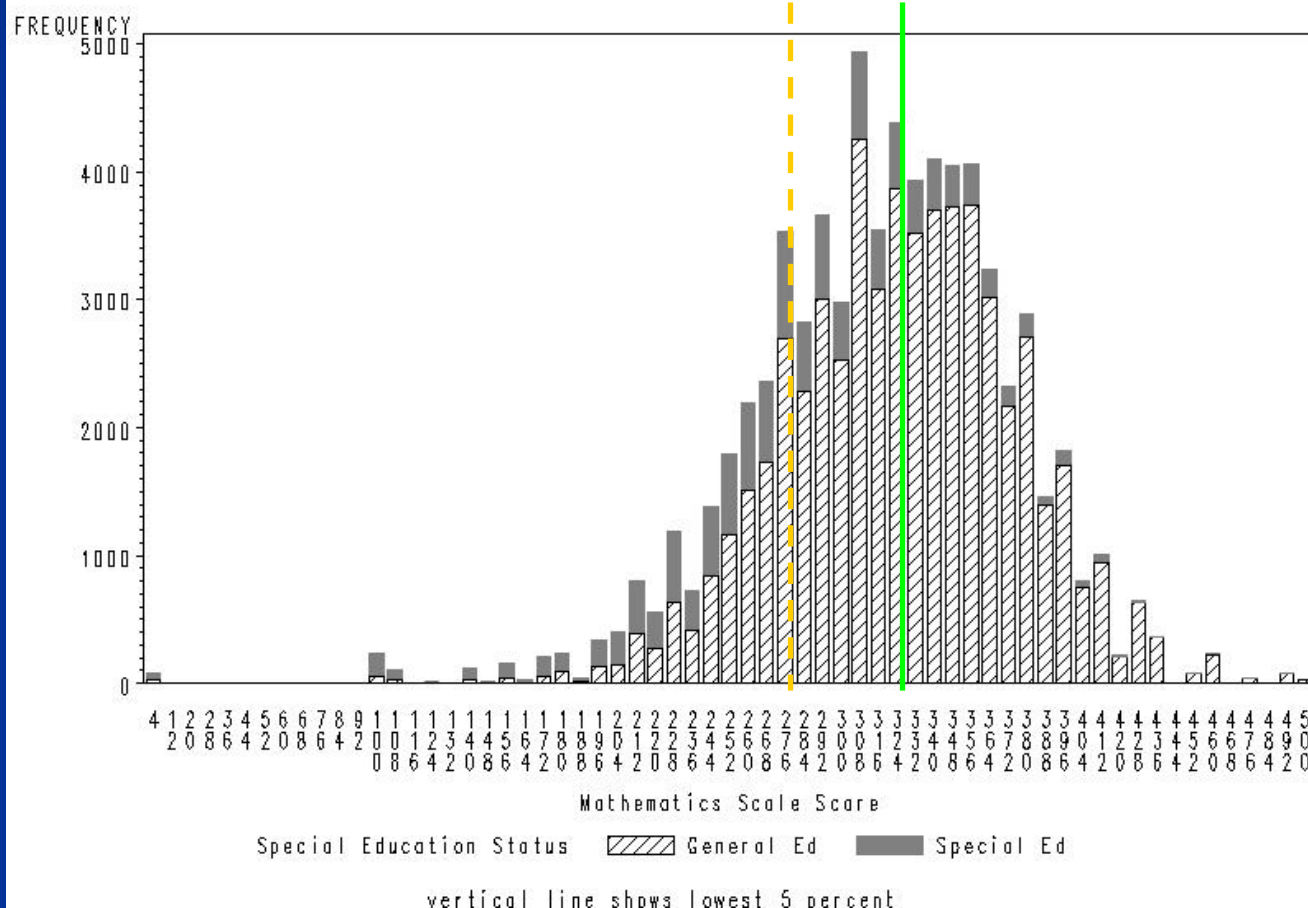
Defining Who Students Are

- Relation to regular content standards, regular achievement level standards
- Relation of students' disabilities to achievement on regular content standards, regular achievement level standards
- Growth towards proficiency on regular assessment (content and achievement levels)



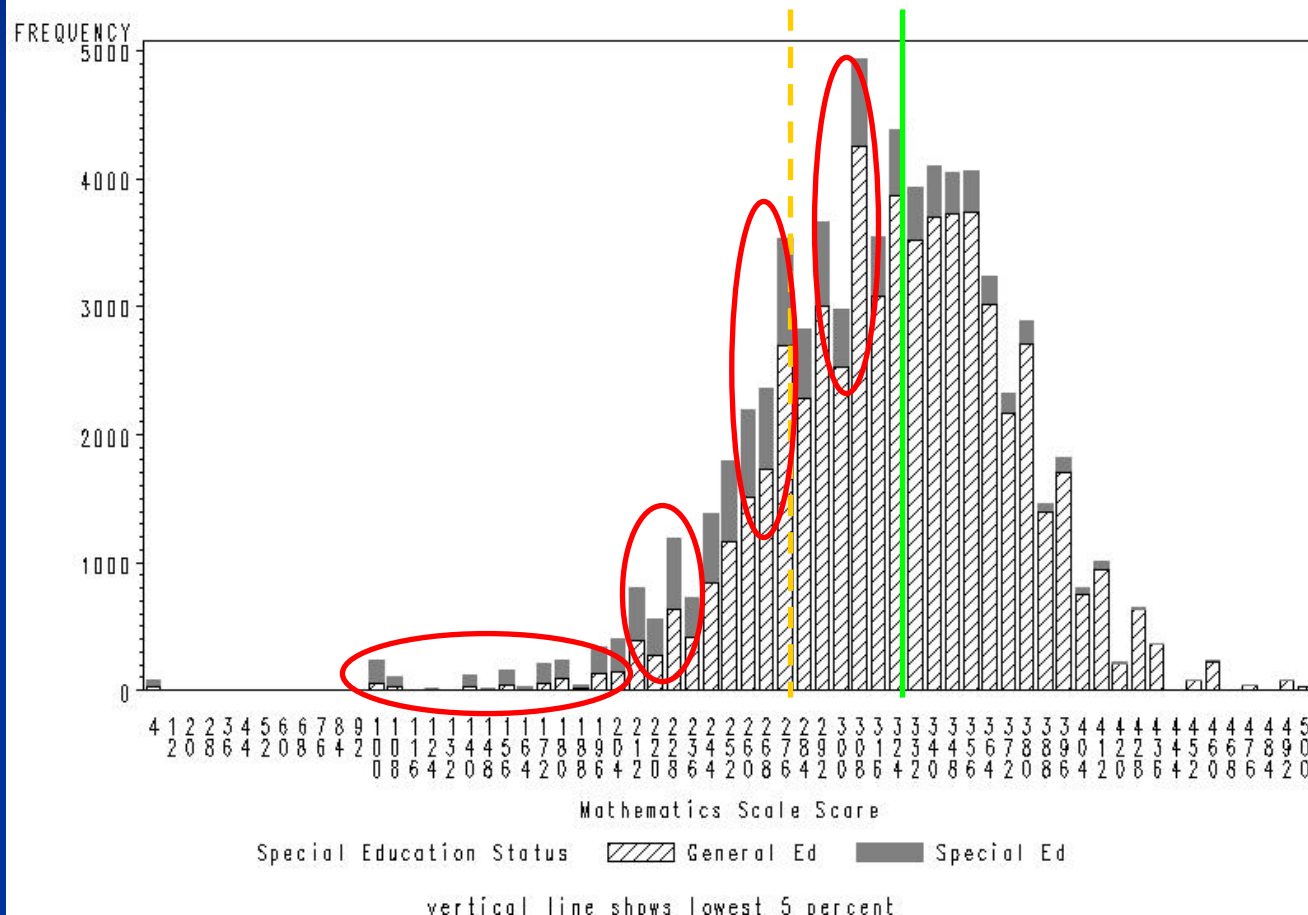
Performance in Relation to Proficient

State 2: Grade 4 Mathematics Scale Scores by Special Education Status



Performance in Relation to Proficient

State 2: Grade 4 Mathematics Scale Scores by Special Education Status



Performance in Relation to Disability Category

Disability	State 1		State 2	
	Mathematics	ELA	Mathematics	ELA
MR	6 %	2 %	19 %	23 %
LD	33 %	33 %	19 %	11 %
ED	36 %	36 %	15 %	13 %
SLI	54 %	55 %	46 %	47 %
MD	--- \$	--- \$	42 %	65 %
HI	30 %	42 %	28 %	19 %
OI	57 %	42 %	30 %	35 %
OHI	41 %	44 %	27 %	21 %
VI	--- \$	--- \$	34 %	32 %
AUT	54 %	45 %	44 %	53 %
TBI	--- \$	--- \$	26 %	29 %
DB	--- \$	--- \$	--- \$	--- \$
DD	--- \$	--- \$	--- \$	--- \$
Overall- Special Ed	39 %	39 %	26 %	22 %
Overall- General Ed	71 %	74 %	62 %	62 %

S - n < 10



Growth Towards Proficiency

Year 1 Performance Level	Year 2 Performance Level				
	I	II	III	IV	V
I	64	27	8	0	0
II	24	43	32	1	0
III	4	18	64	13	1
IV	0	2	39	51	8
V	0	0	10	53	37

Source: Hill, R. Using value tables to explicitly value growth. Maple Grove, MN: JAM Press.



Normative Growth

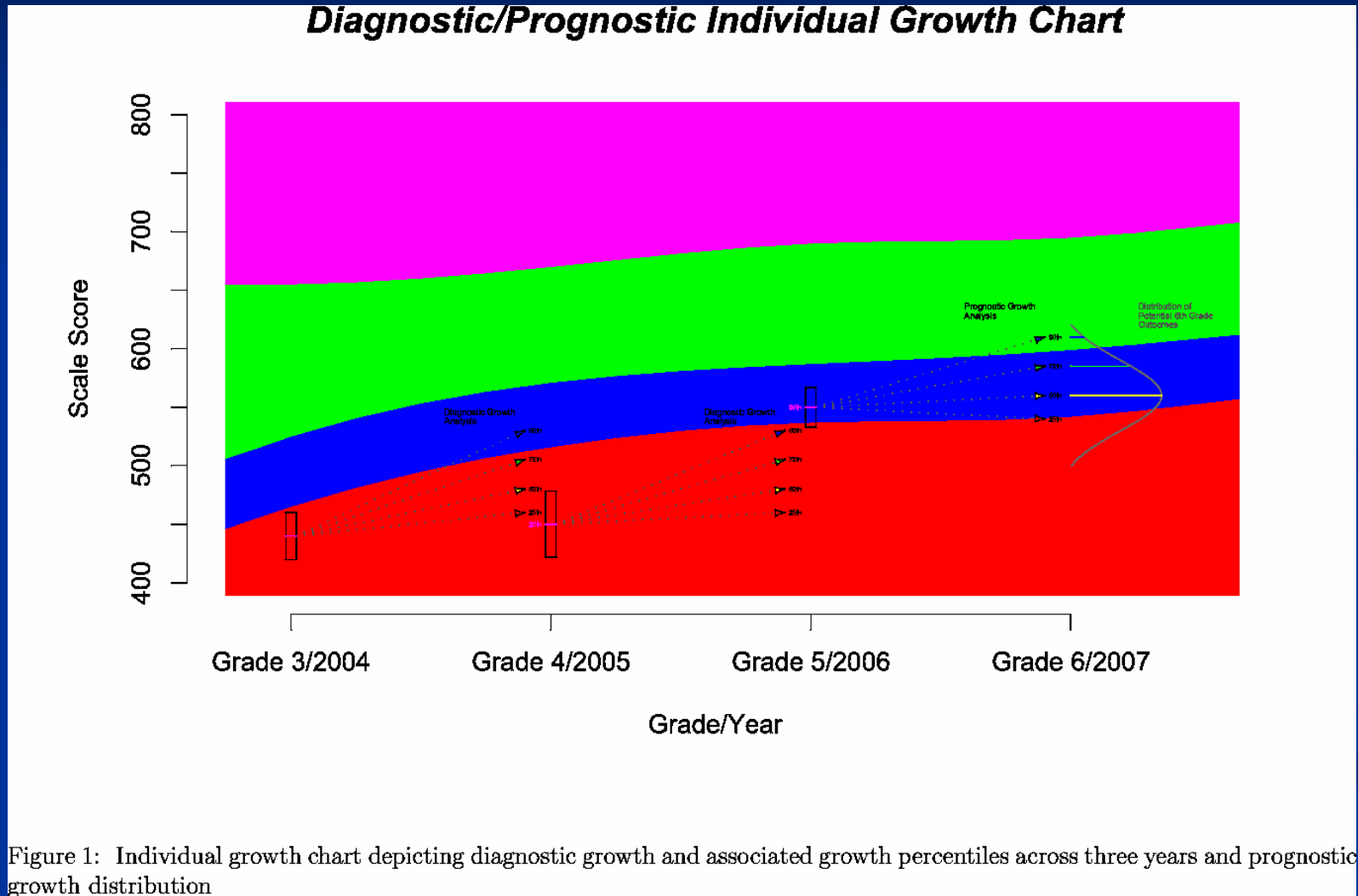
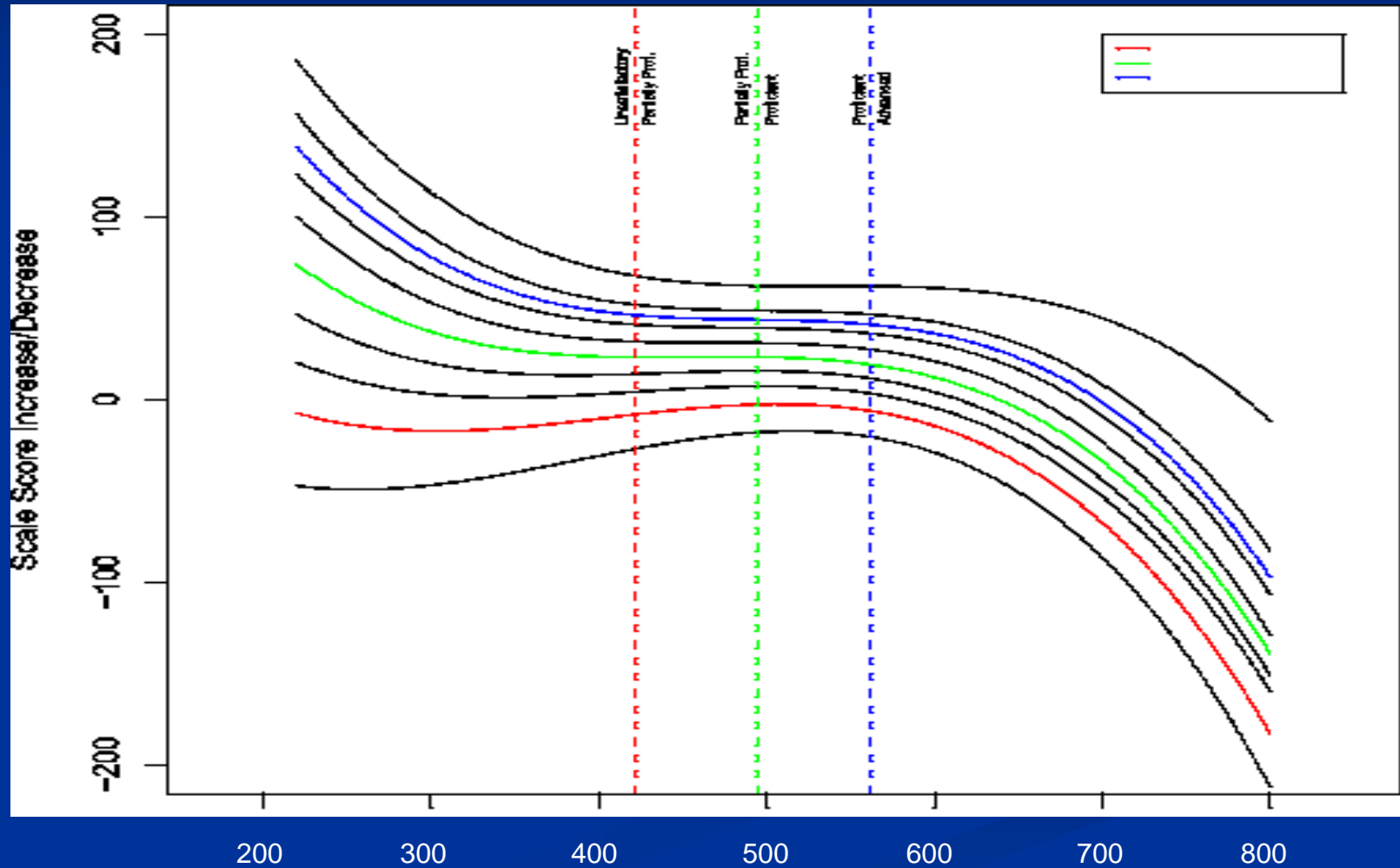


Figure 1: Individual growth chart depicting diagnostic growth and associated growth percentiles across three years and prognostic growth distribution



Normative Growth- 2



How Could Students Grow?

- Towards Proficient
 - Number of years? – 2? 3? 4? 5? 6? ...?
- Towards Proficient-M (MAS)
 - Number of years? – 0? 1? 2? 3? 4? 5?
- If a student grew from starting point to Proficient-M, how long would it be expected to grow to Proficient-R?
- Normative
- Keep up with regular curriculum progression (promotion year-by-year) but not Proficient (at least by 2013-14)



Defining the Construct/Domain

- Assumption is that student is being instructed in the general curriculum
- What is different?
 - Less extensive mastery of same terminal content
 - Working on precursor content that would fall in within-grade instruction for regular curriculum
 - Working on less generalized applications
 - Working in more structured problems or less independent settings



Less Extensive Mastery of Same Terminal Content

- Less extensive mastery of same terminal content
 - Grade 7 mathematics, content strand: Data, statistics, and probability; topic: measures of central tendency and dispersion. Content standard: Students will demonstrate understanding of measures of central tendency. They will be able to construct or calculate mean, median, and mode, given a set of data. They will be able to interpret each type of measure and explain when and why one measure is more appropriate. Assessment limits: data may be any widely used unit (e.g., whole number, dollar amounts, commonly occurring categorical items). The data set will have no more than 15 items.
 - 2% content standard: Student will be able to construct a graph of a given data set and identify the mode. The student may be provided support in constructing the graph, including setting up the axes, simplifying the unit axis to “less than, equal, or more than.”



Less Extensive Mastery – Student Example



How students learn the content: All seventh-graders are learning the concepts of mean, median and mode. They plot various sets of data, including prices, to illustrate the concepts. Ron is plotting the mode using prices cut from advertisements and then glued on an organizer to create a bar graph.

Why this is useful: Looking at information and drawing conclusions from it (data analysis) is an important skill that helps us understand everything from shopping to social trends.

Combining academic and function learning: Ron is learning the concept of more, equal ("same"), and less in the context of consumer choices. Having access to the same information as other students his age helps him develop appropriate language and provides increased opportunities for interaction and communication.

Source: Reformatted from U.S. Department of Education, Office of Special Education and Rehabilitative Services. (2007). *Learning Opportunities for Your Child Through Alternate Assessments*, p. 8 (reformatted). Washington, DC: Author.



What is Different in this Example

- On grade level content – Yes
- Less extensive mastery of same terminal content - Yes
- Working on precursor content that would fall in within-grade instruction for regular curriculum – Yes and No (“More, Less, Equal to” not on grade level)
- Working on less generalized applications - Yes
- Working in more structured problems or less independent settings – Yes



Construct/Domain

- Is that performance target OK with your construct?
- This example showed some differences for one content standard. What about your collection of content standards?
 - What is the total difference over all your standards?



Precursor (Grade Level) Content

- Working on precursor content that would fall in within-grade instruction for regular curriculum (e.g., Finding the fractional part of a whole)
 - Understanding the difference between the whole and the part
 - Partitioning an area into an even number of parts (2, 4, 8, 16...), such as “fair share” problems – how can 4 people share one brownie?
 - Partitioning an area into an odd number parts
 - Identifying the fraction described by the partitioned area
 - Using multiplicative reasoning (division and multiplication) – If students don’t use multiplicative reasoning, then they will always think about fractions like $\frac{3}{4}$ as 3 out of 4, instead 3 out of four *equal parts*
 - Finding a fractional part of a set or area where the number of objects in the set or the number of parts in the area is a multiple of the magnitude of the denominator (e.g., What is $\frac{3}{4}$ of a set of 16 objects?)



Defining Proficiency

- Progress-based?
- Individual for student reflecting goals? – like S-IEP
- Common (Growth)—based on set amount of progress or relation of progress to status goal?
- Common (Status-R)—status-based, related to regular assessment and PLDs (e.g., anchored to Basic in regular assessment)
- Common (Status-M)—status-based, MAS – defined carefully and specifically



Defining Proficiency – 2

- Proficiency only defined (well) in relation to
 - Construct/Domain and
 - Who Students Are and Purpose AND
 - Instructional Model
 - How to make compatible: growth for identification, individual growth for instruction, but Status for AYP



Defining reports

- Scores, Achievement Levels, Inferences & Uses



Defining Assessment Specifications

- Domain sampling and representation (content and performance with content)
 - Reliable
 - Generalization
 - Application
 - Independence
- Assessment limits
- Item level and form (test level) e.g., How many items per content standard?



Some Development Approaches

- Reduce number of standards addressed
- Reduce number of items per standard on test
- Use only “easy” items
- Revise to make items easier – different than regular assessment
- What are you assessing – what is the construct?



Reduce Number of Test Items

- Are you reducing
 - A) Scope
 - B) Generalizability
 - C) Reliability of student-level measurement (status and growth)
 - D) All of the above

- How is that reflected in your construct, definition of proficiency (ALDs), and test specifications?
- How is that related to your purpose/instructional purpose?



Use only “Easy Items”

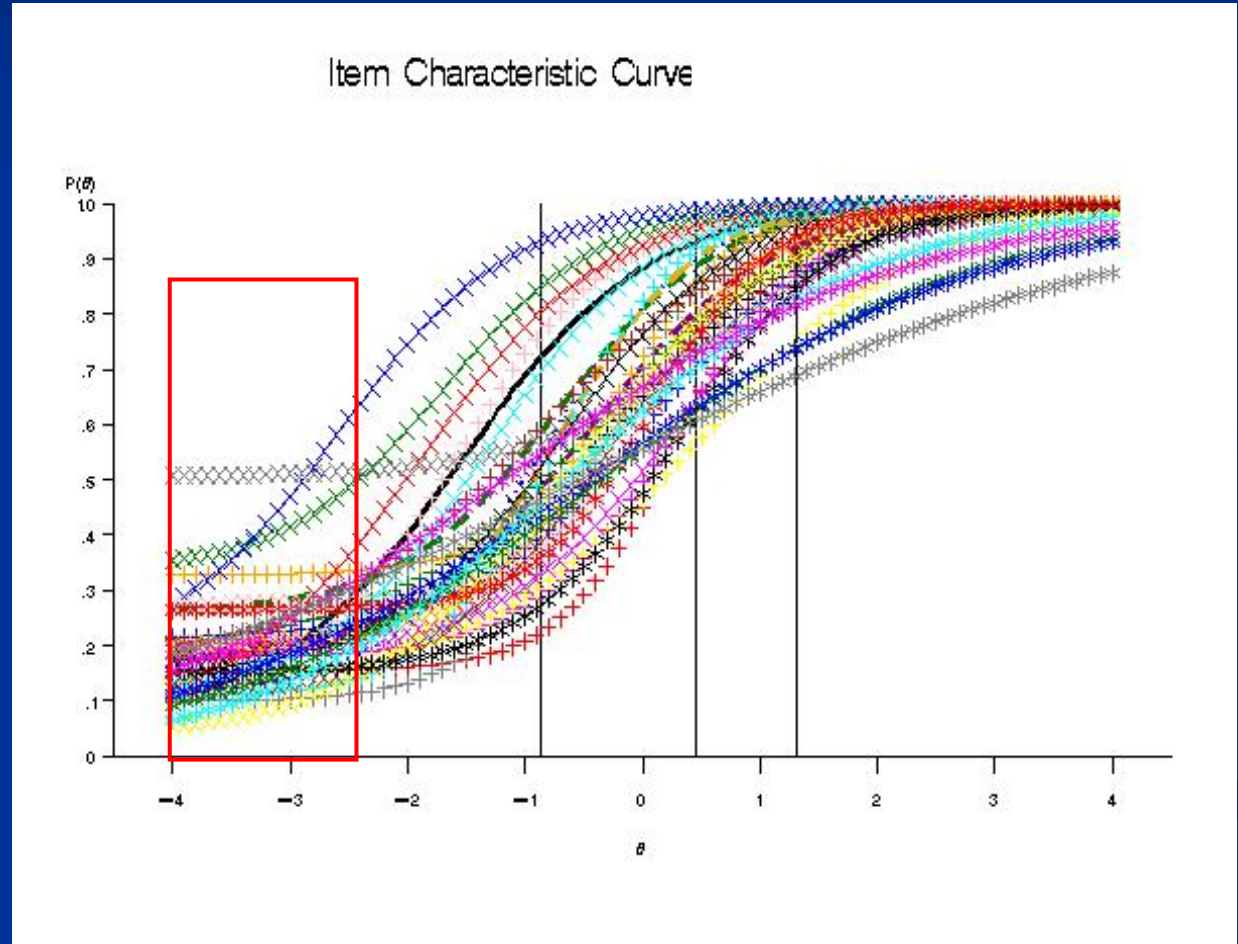
- Item bank with many items addressing all content standards

Ability/Difficulty (Theta)	+4																			
	+3																			
	+2																			
	+1																			
	0																			
	-1																			
	-2																			
	-3																			
	-4																			
			Content Standards (A-N)																	



Check “Easy Items”

- Number
- Information
- Reliability



Use only “Easy” Items

- Are you reducing
 - A) Scope
 - B) Generalizability
 - C) Reliability of student-level measurement (status and growth)
 - D) All of the above

- How is that reflected in your construct, definition of proficiency (ALDs), and test specifications?
- How is that related to your purpose/instructional purpose?



Revise Items to Make “Easier”

- Reduce number of choices in multiple-choice

A)	60 *	A)	63	A)	80
B)	5	B)	---	B)	6
C)	25	C)	27	C)	
D)	10	D)	10	D)	14



Revised “Easier” Items

- Are you reducing
 - A) Scope
 - B) Generalizability
 - C) Reliability of student-level measurement (status and growth)
 - D) All of the above

- How is that reflected in your construct, definition of proficiency (ALDs), and test specifications?
- How is that related to your purpose/instructional purpose?



Considering Technical Aspects

- Equating...



Considering Non-technical Aspects

- Time lines
 - Specs and RFP by early 2008; pilot in spring 2009; administer & report for AYP in spring 2010
- Support resources
 - Item development
 - Equating
 - Administration
 - Training in field
- Budgets



Putting the Pieces Together

Purpose	Who Students Are
Construct/Domain	Proficiency
Reports (inferences, uses)	Assessment Specifications

Instruction (and placement)

- Policy clarification & high quality implementation



For more information:

Center for Assessment

www.nciea.org



Brian Gong

bgong@nciea.org

