IES Practice Guide - Response to Intervention in Mathematics

A Focus on Assessment

Ben Clarke, Ph.D.
Research Associate - Center for Teaching and Learning, University of Oregon

Assisting Students Struggling with Mathematics: Response to Intervention for Elementary and Middle Schools

The report is available on the IES website:

http://ies.ed.gov/ncee & http://ies.ed.gov/ncee/wwc/publications/practiceguides/

Why the Guide?

- Increasing recognition of the importance of mathematical knowledge
  - "For people to participate fully in society, they must know basic mathematics. Citizens who cannot reason mathematically are cut off from whole realms of human endeavor. Innumeracy deprives them not only of opportunity but also of competence in everyday tasks". (Adding it Up, 2001)

High Level of Interest in Mathematics Achievement

- National Research Council: Adding it Up
- National Council Teachers of Mathematics: Focal Points
- National Mathematics Advisory Report

State of Mathematics

- Achievement on the NAEP trending upward for 4th/8th grade and steady for 12th grade
  - Large numbers of students still lacking proficient skills
  - Persistent income and ethnicity gaps
  - Drop in achievement at the time algebra instruction begins
- TIMS data indicate significant lower levels of achievement between US and other nations
  - Gap increase over time
- Jobs requiring intensive mathematics and science knowledge will outpace job growth 3:1 (STEM) and everyday work will require greater mathematical understanding

Panelists

- Russell Gersten (Chair)
- Sybilla Beckman
- Ben Clarke
- Anne Foegen
- Laurel Marsh
- Jon R. Star
- Bradley Witzel

Search for Coherence

Panel works to develop 5 to 10 assertions that are:

- Forceful and useful
- And COHERENT
- Do not encompass all things for all people
- Do not read like a book chapter or article

Challenges for the panel:

- State of math research
- Distinguishing between tiers of support

Jump start the process by using individuals with topical expertise and complementary views

Structure of the Practice Guide

- Recommendations
- Levels of Evidence
- How to carry out the recommendations
- Potential Roadblocks & Suggestions

Evidence Rating

- Each recommendation receives a rating based on the strength of the research evidence.
  - Strong
  - Moderate
  - Low

Recommendation 1

Screen all students to identify those at risk for potential mathematics difficulties and provide interventions to students identified as at risk.

- Level of Evidence: **Moderate**

What is Assessment?

Def:

Assessment is the collection of data to make decisions.

(Salvia & Ysseldyke, 1997)

Assessment is useless if we don't use it to guide our actions.

Assessment for Different Purposes

- An effective, comprehensive mathematics assessment program includes assessments for four purposes:
  - Outcome
  - Screening
  - Progress Monitoring
  - Diagnostic

Outcome Assessment

- Purpose: To determine level of proficiency in relation to norm or criterion.
- When: Typically administered at end of year. Can be administered pre/post to assess overall growth.
- Who: All students
- Relation to instruction: Provides index of overall efficacy but limited timely instructional information.

Screening Assessment

- Purpose: To determine children who are likely to require additional instructional support (predictive validity).
- When: Early in the academic year or when new students enter school. May be repeated in the Winter and Spring.
- Who: All students
- Relation to instruction: Most valuable when used to identify children who may need further assessment or additional instructional *support*.

Progress Monitoring Assessment

- Purpose: Frequent, timely measures to determine whether students are learning enough of critical skills.
- When: Weekly or Monthly
- Who: At-risk students
- Relation to Instruction: Indicates student response to instruction.

Diagnostic Assessment

- Purpose: To provide specific information on skills and strategy needs of individual students.
- When: Following screening or at points during the year when students are not making adequate progress.
- Who: Selected students as indicated by screening or progress monitoring measures or teacher judgment.
- Relation to Instruction: Provided specific information on *target skills*; highly relevant.

Coherent Assessment Systems

- Each type of assessment has a purpose
- The design of the tool should match the purpose
  - What are the implications for screening tools used with all students?
- Think purpose not tool
- How do each of these purposes fit together?

Does your school collect data to make decisions or to collect data?

- Common pitfalls
  - Focus is on procedure
  - Data collected don't match purpose for collecting data (e.g. collecting diagnostic data on all students)
  - Layering of data sources
  - Different data for different programs (e.g. Title 1)

Activity: Do your current assessments function as a whole?

- Talk with a colleague about how the four types of assessments work in one system at your school/district.
  - Does each assessment tool match the purpose it is used for?
  - Does the system link together in a logical manner?

Recommendation 1

Screen all students to identify those at risk for potential mathematics difficulties and provide interventions to students identified as at risk.

- Level of Evidence: **Moderate**

Technical Evidence

- Correlational design studies
  - Greater evidence in the earlier grades
  - Reliability typically included inter-tester, internal consistency, test-retest, and alternate form
    - Most fall between r=.8 to .9
  - Validity primarily focused on criterion related with an emphasis on predictive validity
    - Most fall between r=.5 to .7
  - Measures are beginning to report on sensitivity and specificity

Content

- Content of Measures
  - Single aspect of number sense (e.g. strategic counting) – most common in earlier grades
  - Or Broad measures incorporating multiple aspects of number
    - Some measures are combination scores from multiple single aspect measures
  - Measures reflecting the computation and concepts and applications objectives for a specific grade level – most common later grades
    - Often referred to as CBM or General Outcome

Content

- Promising measures include
  - Word problems
  - Pre-algebra and algebra skills
  - Based on state standards or NCTM/NMAP benchmarks

Features

- Short duration measures (1 minute fluency measures)
  - Note many measures that are short duration also used in progress monitoring.
- Longer duration measures (untimed up to 20 minutes) often examine multiple aspects of number sense
  - Issue of purpose is critical to examine
- Most research examines predictive validity from Fall to Spring.

Examples: Single aspect number sense

- Example: Magnitude comparison

12	3	4	1	5	11	9	4

- Example: Strategic counting

__	13	14	6	__	8	3	4	__

Example: Multiple aspects number sense

- Number Knowledge Test
  - Level 1
    - If you had 4 chocolates and someone gave you 3 more, how many chocolates would you have?
    - Which is bigger: 5 or 4?
  - Level 3
    - What number comes 9 after 999?
    - Which difference is smaller: the difference between 48 and 36 or the difference between 84 and 73?

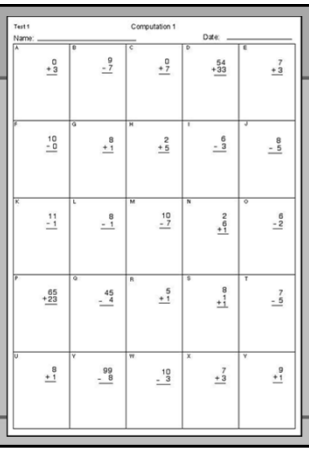
2nd grade and above: Examples

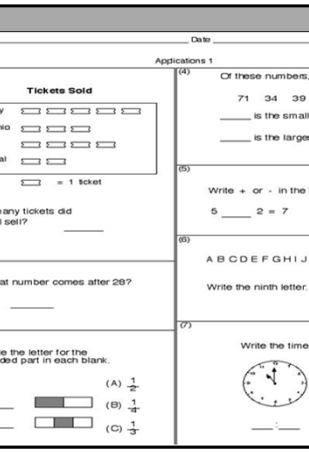
- Number combinations
- Word problems
- Grade level computation objectives
- Grade level concepts and applications
- Measures tied to CCSS; NMAP; Focal Points

General Outcome: Computation and Concepts and Application objectives

- For students in grades 1–6.
- Student is presented with 25 computation or concepts and applications problems representing the year-long, grade-level math curriculum.
- Student works for set amount of time (time limit varies for each grade).
- Teacher grades test after student finishes.

28

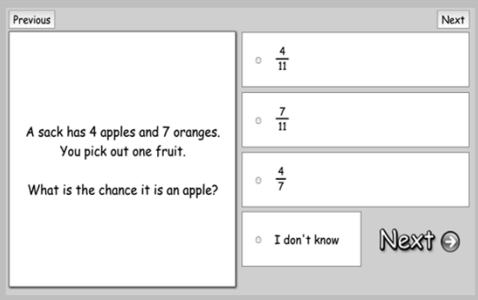




Example: Reflecting critical math content

- easy-CBM
- Items created according to NCTM Focal Points for grade level
- 48 items for screening (16 per focal point)
- Ongoing research (not reviewed in practice guide)

easy-CBM: Number and Operations



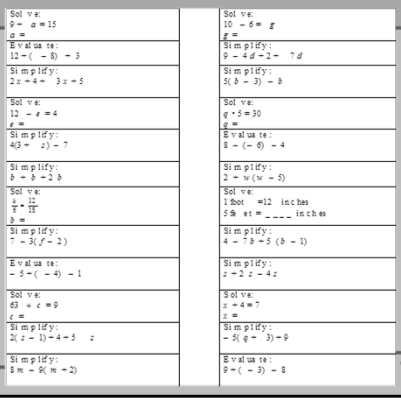
Middle School

- Algebra measures
  - Designed by Foegen and colleagues assess pre-algebra and basic algebra skills. Administered and scored similar to Math-CBM
- Math CBM Computation and Concepts and Applications
  - Concepts and Applications showed greater valdity in 6th, 7th, and 8th grade

Basic Skills (in Algebra)

- 60 items; 5 minutes
- Problems include:
  - Solving basic fact equations;
  - Applying the distributive property;
  - Working with integers;
  - Combining like terms;
  - Simplifying expressions;
  - Applying proportional reasoning
- Scoring: # of problems correct

Basic Pre-algebra skills



Solve: $9 + a = 15$ $a =$	Solve: $10 - 6 = g$ $g =$
Evaluate: $12 \div (-8) + 3$	Simplify: $9 - 4d + 2 = 7d$
Simplify: $2x + 4 = 3x + 5$	Simplify: $5(b - 3) - b$
Solve: $12 - e = 4$ $e =$	Solve: $q \cdot 5 = 30$ $q =$
Simplify: $4(3 + z) - 7$	Evaluate: $8 - (-6) - 4$
Simplify: $b + b + 2b$	Simplify: $2 + w(w - 5)$
Solve: $\frac{b}{3} = \frac{12}{18}$ $b =$	Solve: 1 foot = 12 inches 5 feet = ____ inches
Simplify: $7 - 3(f - 2)$	Simplify: $4 - 7b + 5(b - 1)$
Evaluate: $-5 \div (-4) - 1$	Simplify: $z + 2z - 4z$
Solve: $63 \div c = 9$ $c =$	Solve: $z + 4 = 7$ $z =$
Simplify: $2(z - 1) + 4 + 5z$	Simplify: $-5(q + 3) + 9$
Simplify: $8m - 9(m + 2)$	Evaluate: $9 \div (-3) - 8$

Suggestions

- Have a building level team select measures based on critical criteria such as reliability, validity and efficiency.
  - Team should have measurement expertise (e.g. school psychologist) and mathematics (e.g. math specialist)
  - Set up a screening to occur twice a year (Fall and Winter)
  - Be aware of students who fall near the cut scores

Suggestions

- Select screening measures based on the content they cover with a emphasis on critical instructional objectives for each grade level.
  - Lower elementary: Whole Number
  - Upper elementary: Rational Number
  - Across grades: Computational Fluency (hallmark of MLD)

Suggestions

- In grades 4-8, use screening measures in combination with state testing data.
  - Use state testing data from the previous year as the first cut in a screening system.
  - Can then use a screening measure with a reduced pool of students or a more diagnostic measure linked to the intervention program for a second cut.

Suggestions

- Use the same screening tool across a district to enable analyzing results across schools
  - Districts may use results to determine the effectiveness of district initiatives.
  - May also be used to determine systematic areas of weakness and provide support in that area (e.g. fractions)

Roadblocks

- Resistance may be encountered in allocating time resources to the collection of screening data.
- Suggested Approach: Use data collection teams to streamline the data collection and analysis process.

Roadblocks

- Questions may arise about testing students who are “doing fine”.
- Suggested Approach: Screening all students allows the school or district to evaluate the impact of instructional approaches
  - Screening all students creates a distribution of performance allowing the identification of at-risk students

Roadblocks

- Screening may identify students as at-risk who do not need services and miss students who do.
- Suggested Approach: Schools should frequently examine the sensitivity and specificity of screening measures to ensure a proper balance and accurate decisions about student risk status.

Sensitivity and Specificity

		Students at-risk	
		YES	NO
Students identified as at-risk	YES	True positive (A)	False positive (B)
	NO	False negative (C)	True negative (D)

Sensitivity: Number of students correctly identified as at-risk or A/(A+C)

Specificity: Number of student correctly identified as not at risk or D/(D+B)

Sensitivity and Specificity

- Cut score is set too high:
  - You have good sensitivity (all kids that need help are identified) but poor specificity (lots of kids who don't need help are identified)
- Cut score is set too low:
  - You have good specificity (most kids who don't need help will not be identified as at-risk) but poor sensitivity (you may miss many kids who do need help)

An example - easyCBM

- Sensitivity at least .90 - Johnson, Jenkins, Petscher, & Catts (2009)
  - Favors higher cut scores
- Sensitivity and Specificity at least .70 - Silberglitt & Hintze (2005)
  - Favors lower cut scores

Slide 43

BC1 Ben Clarke, 10/2/2009

Discussion cont.

- Winter 25th%ile criteria
  - Johnson procedure = cut of 34
    - 70 students identified as at-risk
    - 22 truly at-risk
    - 48 false positives (provided non needed services)
    - 1 false negative (not provided needed services)

Discussion cont.

- Winter 25th%ile criteria
  - Silberglitt procedure = cut score 30
    - 41 students identified at-risk
    - 18 truly at-risk
    - 23 false positives
    - 5 false negatives

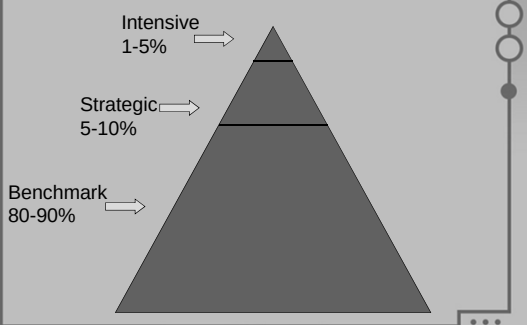
Example cont.

- To identify 4 additional at-risk students; you over identify an additional 29 students
  - If small group instruction provided (3-5 students per group) an additional 6-10 groups are needed.
  - Impact on limited school resources
- Schools rarely discuss what “at-risk” means

Roadblocks

- Screening may identify large numbers of students who need support beyond the current resources of the school or district.
- Suggested Approach: Schools and districts should
  - Allocate resources to the students with the most risk and at critical grade levels
    and
  - Implement school wide interventions to all students in areas of school wide low performance (e.g. Fractions)

In a Three-Tier Model…



But in reality….

2009 NAEP Performance Data - National Level

	Grade 4	Grade 8
All Students	39%	34%*
SWD	19%	9%
ELL	12%	5%
Eligible for F/R lunch	22%	17%

Roadblocks

- Screening may identify large numbers of students who need support beyond the current resources of the school or district.
- Suggested Approach: Schools and districts should
  - Allocate resources to the students with the most risk and at critical grade levels
    and
  - Implement school wide interventions to all students in areas of school wide low performance (e.g. Fractions)

Activity: What is your current screening process?

- What measure(s) do you use?
- What works well in your system?
- What roadblocks have you encountered?

Recommendation 7

- Monitor the progress of students receiving supplemental instruction and other students who are at risk
  - Level of evidence: Low

Evidence

- Non-experimental studies demonstrating the technical adequacy of progress monitoring measures.
  - Reliability and Validity are similar to that found for screening measures (often the same measure)
  - Growth has been typically examined by looking at average scores across time
  - Some evidence of use in instructional decision making and improved student outcomes
- Greater evidence in elementary grades

Content and Features

- General outcome measures reflecting concepts and computation objectives for the grade level.
  - Some limited evidence for single aspect measures (i.e. Magnitude comparison)
- All are timed and short duration

Suggestions

- Monitor the progress of tier 2, tier 3 and borderline tier 1 students at least once a month using grade appropriate general outcome measures.
  - Same team that worked on screening can also work on progress monitoring
  - Need to carefully consider capacity to model growth in the context of instructional decision making

Suggestions

- Use curriculum-embedded assessments in intervention materials
  - Frequency of measures can vary - every day to once every week.
  - Will provide a more accurate index of whether or not the student is obtaining instructional objectives
  - Combined with progress monitoring provides a proximal and distal measuue of performance

Roadblocks

- Students within classes are at very different levels.
- Suggested Approach: Group students across classes to create groups with similar needs.

Roadblocks

- Insufficient time for teachers to implement progress monitoring.
- Suggested Approach: Train paraprofessionals or other school staff to administer progress monitoring measures.

Activity: What is your current progress monitoring process?

- What measure(s) do you use?
- What works well in your system?
- What roadblocks have you encountered?

How to start and next steps

- Focus on one grade or grade bands
- Accumulating evidence that math trajectories are established early and difficult to alter
  - Students entering and exiting kindergarten with poor mathematics achievement (defined as below the 10th percentile at both times on a nationally normed mathematics assessment) had a 70% chance of scoring below the 10th percentile 5 years later (Morgan et al., 2009)

How to start and next steps

- Staring early (K-1) may be smart and strategic
  - Greater comfort with whole number content and instruction
  - Greater array of research based instructional programs

How to start and next steps

- Screening before progress monitoring
- Strategies for collecting data
- Open dialogue

Planning for this year

- Take 5 minutes to debrief on content
- Take additional time to plan or ask additional question

THANKS!