Benchmark Data Consult
"The Sort…"

From Beth Wood
To Randy Randolph
In preparation for October 17

Types of Assessment

- Within schools we must organize our assessment "toolkits" around <u>four</u> broad types of assessment instruments
  1. Screening assessments
  2. Diagnostic assessments
  3. Progress Monitoring assessments
  4. Mastery measures (*too little too late for our purposes today)*
- **As educators we must understand each assessment's <u>purpose</u> and know how to <u>use</u> the resulting data**

Beth Wood, 2011

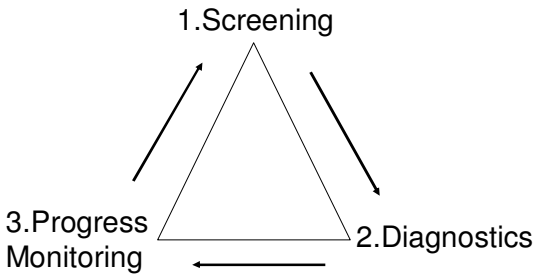
Data Based Decision-Making

- An adequate *formative assessment plan* is imperative
  - Screening tools
  - Diagnostic tools
  - Progress Monitor tools

- Select norms for consistent use
- Establish benchmarks/cut scores
- Teach staff to
  - Collect baseline data
  - Construct appropriate goal lines
  - Read student trend lines and make decisions

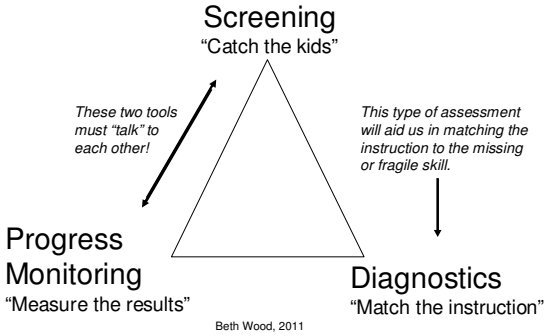
Beth Wood, 2011

Triangulate the Data



Beth Wood, 2011

Triangulate the Data:
Catch, Match, & Measure



Beth Wood, 2011

Curriculum Based Measures

- Are scientifically valid and reliable for screening and progress monitoring
- Are single-trial-effective (with the exception of the R-CBM; administer 3 and use the median score)
- Assess two things by measuring *fluency*
  - Accuracy (and this should always be the 1st consideration)
  - Rate of response (effortlessness, automaticity, fragility)
- Are not *diagnostic*…by design
- Are timed for two purposes
  - To check for *normal* rate of response (faster is NOT better)
  - For teacher *utility* (conserving instructional time)

Beth Wood, 2011

Do's and Don'ts

- **Do**
- Begin with the end in mind…
- Use the Benchmark to identify the major skill areas where students are struggling
- Use *diagnostics…* to make the instructional match- - -meaning prerequisite sub-skills

- **Don't**
- Use CBM probes to *practice* the test to just increase a student's score
- Use CBMs for any purpose other than what they were designed for: screening and progress monitoring

Beth Wood, 2011

Math RtI/MTSS

- Math is handled in much the same way Reading is treated in a tiered system
- Screen to determine who is at risk for Math failure
- There may not be the diagnostic tools available, so we rely on "error analysis" to determine where to begin the instruction (increasing skill complexity)
- Progress Monitor Math students just as one would Reading students
- Set realistic and ambitious goals and use the data to inform further supplemental instruction

Beth Wood, 2011

Outcomes

- A Benchmark Universal Screening will tell us if students can demonstrate an outcome:
  - Name letters or associate sounds
  - Segment/blend words or non-words
  - Read fluently and with prosody
  - Comprehend *syntactically (sentence level)*

  symptoms
- When a score indicates that a student cannot perform an outcome skill proficiently, we can go to a diagnostic tool in our toolkit to learn what skill or skills in the sequence are missing or fragile; in other words… where the learning left off.

Beth Wood, 2011

causes…

Skill Deficit or Fragile Skill

- Require different types of intervention
- Lack of accuracy indicates a skill **deficit**
- A skill deficit (or missing skill) requires modeling, direct and explicit instruction, and practice
- An accurate but slow rate of response indicates the student can perform the skill with much effort…and time…it's **fragile**
- A fragile skill requires aligned and sufficient amounts of massed and distributive practice to promote effortlessness (work together to help students generalize)

Beth Wood, 2011

The Benchmark

- Administer each *recommended* measure during the official fall, winter, and spring 'window' (late arrival new students)
- Train personnel who will be administering each measure; plan for fidelity checks
- Administer each Benchmark measure using the standardized method
  - Scripted directions, timing, test environment, etc.
- Use standardized scoring procedures to ensure valid and reliable data results

Beth Wood, 2011

The Benchmark

- Carefully enter each student's data into the system
- Enter both "corrects" and "errors" if appropriate
- Personnel should be aware of previously agreed upon benchmark "cut scores" developed to identify students at risk for reading and/or math failure
- Decisions should also be made to determine at what point students will be Survey Level Assessed to determine appropriate progress monitoring levels
- Appropriate personnel should run grade level or class Distribution Reports (in black and white) for use during the Benchmark Data Consult and 'sort for instruction'

Beth Wood, 2011

Data Consult

- Each fall, winter and spring, use grade or class level distribution reports, to identify students who scored below the benchmark.
- Use previously agreed upon cut scores for each measure (from multi-year aggregate norm tables)

Choose your cut score

Back | E-Mail | PDF

Below Above Above & Below Target
Show Grade Show Target
View Chart

Class Distribution by Scores and Percentile
Red River Valley District - Jefferson Elementary
Grade 3 - (Date Range)(Spring 2003-2004)
Reading - Curriculum Based Measurement

Name	Correct	Errors	Accuracy	Performance Summary	Potential Instructional Action
Hudson, Greg	211	1	99.5%	Well Above Average	Consider Need for Individualized Instruction
Schumacher, Haris	196	3	98.5%	Well Above Average	Consider Need for Individualized Instruction
Clarkine, Lindsay	195	3	98.5%	Well Above Average	Consider Need for Individualized Instruction
Fenger, Sydney	194	1	99.5%	Well Above Average	Consider Need for Individualized Instruction
Gehman, Kianna	180	6	96.8%	Well Above Average	Consider Need for Individualized Instruction
Well Above Average >= 180 (90th %ile)					
Clark, Tyler	170	6	96.6%	Above Average	Consider Need for Individualized Instruction
Scanlon, Zachary	167	2	98.8%	Above Average	Consider Need for Individualized Instruction
Jensen, Kevin	161	4	97.6%	Above Average	Consider Need for Individualized Instruction
Sells, Haley	158	0	100.0%	Above Average	Consider Need for Individualized Instruction
Above Average >= 150 (75th %ile)					
Orth, Joey	130	2	98.5%	Average	Continue Current Program
Connaster, Ryan	123	1	99.2%	Average	Continue Current Program
Dunbar, Ellen	122	2	98.4%	Average	Continue Current Program
Mowry, Sandra	118	0	100.0%	Average	Continue Current Program
Odegard, Kristina	117	3	97.5%	Average	Continue Current Program
Reidlow, Derrick	115	1	99.1%	Average	Continue Current Program
Target = 107					
Mackey, Damon	108	1	99.1%	Average	Continue Current Program
Williams, Jessica	104	3	97.2%	Average	Continue Current Program
Bloder, Hailey	89	2	97.8%	Average	Continue Current Program
Arias, Alex	87	4	95.6%	Average	Continue Current Program
Johnz, Carissa	86	1	98.9%	Average	Continue Current Program
Average >= 66 (25th %ile)					
Howring, Kori	70	4	94.6%	Below Average	Further Assess and Consider Individualizing Program
Nordrum-Mirst, Nate	68	1	98.6%	Below Average	Further Assess and Consider Individualizing Program
Below Average >= 64 (10th %ile)					
Shapiro, Lane	67	3	95.0%	Well Below Average	Begin Immediate Problem Solving
Simkins, Rachenal	53	3	94.6%	Well Below Average	Begin Immediate Problem Solving
Hugel, Philip	30	7	81.1%	Well Below Average	Begin Immediate Problem Solving

Fall Winter Spring

Beth Wood, 2011

What do percentiles mean?

- 90th percentile – well above average
- 76th percentile – above average
- 75th percentile – average range (high)
- 50th percentile – average range (median)
- 25th percentile – average range (low)
- 24th percentile – below average range (mild)
- 10th percentile – well below average (severe)

Beth Wood, 2011

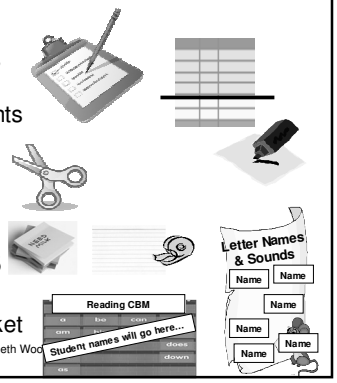
Identifying students

- Run class and/or grade level Distribution Reports for each measure administered to students (in *black&white…save your ink!)
- Using previously agreed upon "cut scores", mark a bold line at the agreed upon *minimum acceptable* score; typically just beneath the 25 %ile *using national norms*
- Students below this bold line would be considered below Benchmark or expectation, at risk for failure, and in need of supplemental instruction

Beth Wood, 2011

Here's what we'll need to sort…

- Class rosters and Distribution Reports (rainbow reports) to be certain all students are accounted for…
- Grade Level and measure cut scores (**norms**)
- Post-It notes or 3X5 note cards
- Chart paper or pocket charts



Beth Wood, 2011

Benchmark Data Consult Sample

- Reports
  - Grade level or classroom distribution reports (by scores and percentiles)
  - Aggregate norms/cut scores (now **stratified**)
- Forms
  - P.A.S.T. & Q.P.S.
  - 4 Quad Analyzer
  - MAZE tool
  - Goal setting form
  - Survey Level Assessment Table

1. Use "rainbow" report to identify students
2. Use *diagnostics* to sort for instruction
3. Group students
4. Plan SLA
5. Set goals
6. Select interventions, implementer, materials
7. Plan progress monitor tool and schedule
8. Plan Fidelity ✓s
9. Update Data Wall

Beth Wood, 2011

"Tagging" students

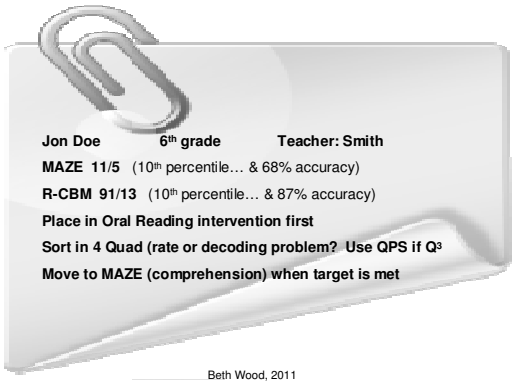
- When a student is below the agreed upon Benchmark cut score, put his or her name on a Post-It note or 3X5 note card
- Note the student's grade level, teacher, **all** measures on which the student did <u>not</u> meet the cut score, their actual scores (35/7), and any follow-up diagnostic tool that will need to be administered to appropriately plan aligned instruction (with skill and score)
  - Alphabet checklist, P.A.S.T., Q.P.S., MAZE 4-Square Tool, any formal diagnostics, etc.

Beth Wood, 2011

The **first** sort…

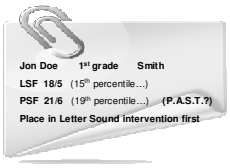
- Using large sheets of chart paper or a number of pocket charts, label sheets or charts with the names of the measures administered to students (grade levels do not need to be on different sheets or charts if you have a school wide block, but may be grouped by grade level)
- Place students names/post-its/3X5 cards on the appropriate measure where they did not meet the cut score
- If the student falls below the cut-score on more than one measure, put his/her name on the *lowest pre-requisite skill,* and that is where instruction will begin for that student (LSF vs. PSF)

Beth Wood, 2011



Beth Wood, 2011

Letter Sound Fluency



Phoneme Segmentation Fluency

Beth Wood, 2011

Now we know **who** they are… but **what** do we <u>teach</u> them?

- When a student scores below the grade level benchmark on a CBM **g**eneral **o**utcome **m**easure, that's telling us they cannot demonstrate a broad outcome like segmenting and blending sounds into words, reading printed words correctly and effortlessly, or comprehending and making adequate meaning from text
- We will have to use something more *diagnostic* to learn what specific skills the student may be deficit in…

Beth Wood, 2011

Diagnostics

- Must help us make an *instructional match*, plan skill-specific instruction targeted at a missing or fragile skill
- Must be do-able; teacher friendly
  - Quick (not necessarily timed)
  - Easy to administer (not necessarily standardized)
  - Easy to score and interpret; very skill-specific and cut score is a minimum number correct out of attempts (8/10)

Beth Wood, 2011

Diagnostics

- Commercial tools
  - DRA, SRI, F&P, Flynt-Cooter, Gates MacGinite, etc.
- Informal tools
  - Checklists to establish patterns (LNF, LSF)
  - P.A.S.T. *Phonemic Awareness Skills Test* (PSF, NWF)
  - QPS *Quick Phonics Screener (R-CBM)*
  - 4 Quad Analyzer (Oral Reading *Fluency)*
  - 4 Square Maze Tool (Comprehension)
  - Available common assessments

Beth Wood, 2011

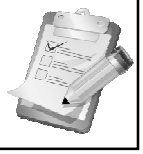
Using Informal Diagnostics to Aid in Sorting for Instruction

- Letter Name and Letter Sound Fluency
  - Use a checklist to look for patterns regarding which letters/sounds are known/unknown (optional)
- Phoneme Segmentation Fluency
  - Phonemic Awareness Skills Test (PAST)
- Reading-CBM (oral reading fluency)
  - 4 Quadrant Sort
  - Quick Phonics Screener (for Q^3)
- MAZE (syntactic comprehension/fluency)
  - 4 Square Tool (topic, main idea, details, summarizing, inferring, etc.)

Beth Wood, 2011

The **next** sort…

- Following the administration and scoring of informal diagnostics, put student scores on their post-it or 3X5 and sort students into small groups (on the chart paper) who have the same or similar sub-skill deficits
- Groups of 4 to 6 are ideal for supplemental reading instruction, but groups may initially be a bit larger
- Try to group students according to where their skills or learning *left off* on the informal diagnostic tool



Beth Wood, 2011

Error Analysis for Math Measures

- When students perform poorly when counting, identifying numbers correctly, knowing which number is larger/smaller, or being able to identify the number that comes between two numbers, the instruction is straight forward
- With computation and concepts and applications, we will need to do error analysis to determine which sub-skills require supplemental instruction; where do skills begin to drop off?

Beth Wood, 2011