Critical Features of Progress Monitoring (with Curriculum-Based Measures)

Gerald Tindal

Castle-McIntosh-Knight Professor

Behavioral Research and Teaching – UO



Three Big Ideas

Three big ideas are presented in this address, covering the essentials of curriculum-based measurement (CBM), requisite features to expect of any progress monitoring system, and critical cautions in using data to make decisions. Although CBM appears to be ubiquitous and applied to almost any type of progress monitoring system, teachers must understand what it is and what it isn't. What should progress monitoring systems look like? How should they be used? What are some of the requisite features and what is the potential for impact? The presentation concludes with cautions in moving forward, if only to ensure success.

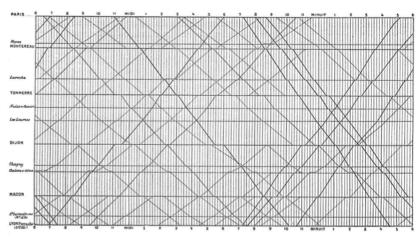


An Overview in Three Parts

Part 1: A history of nearly everything on CBM

Part 2: What a formative assessment should have (with attention to easyCBM)

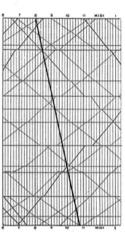
Part 3: Be careful of the question you ask: Is it working?



E. J. Marey, La Méthode Graphique (Paris, 1885), p. 20. The method is attributed to the French engineer, Ibry.

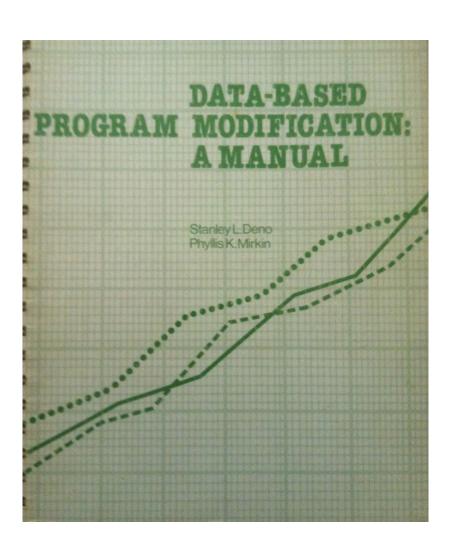
A design with similar strengths is Marey's graphical train schedule for Paris to Lyon in the 1880s. Arrivals and departures from a station are located along the horizontal; length of stop at a station is indicated by the length of the horizontal line. The stations are separated in proportion to their actual distance apart. The slope of the line reflects the speed of the train: the more nearly vertical the line, the faster the train. The intersection of two lines locates the time and place that trains going in opposite directions pass each other.

In 1981 a new express train from Paris to Lyon cut the trip to under three hours, compared to more than nine hours when Marey published the graphical train schedule. The path of the modern TGV (train à grande vitesse) is shown, overlaid on the schedule of 100 years before:

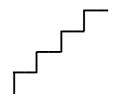




In the beginning...

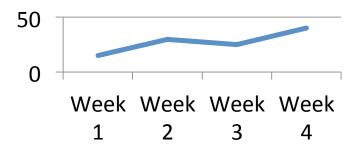


Mastery monitoring: Time to master a set of instructional objectives



General outcome measurement:

Performance graphs reflecting how a student's behavior changes on a single task...over time





Original Conditions and Emphases

- Technically adequate: Must be valid...
- Sensitive to relatively small adjustments made in: instructional methods and materials, motivational techniques, administrative arrangements
- Easy to develop and administer
- Alternate forms available to administer frequently
- Time efficient
- Inexpensive
- Unobtrusive
- Simple to teach



And an atheoretical approach...

"Our hope is that regardless of personal philosophical, theoretical, historical, and current situational constraints, those responsible for ensuring the quality of learning disabilities services will continuously evaluate the impact of those services on the academic and social behavioral of their individual students" (p. 4).

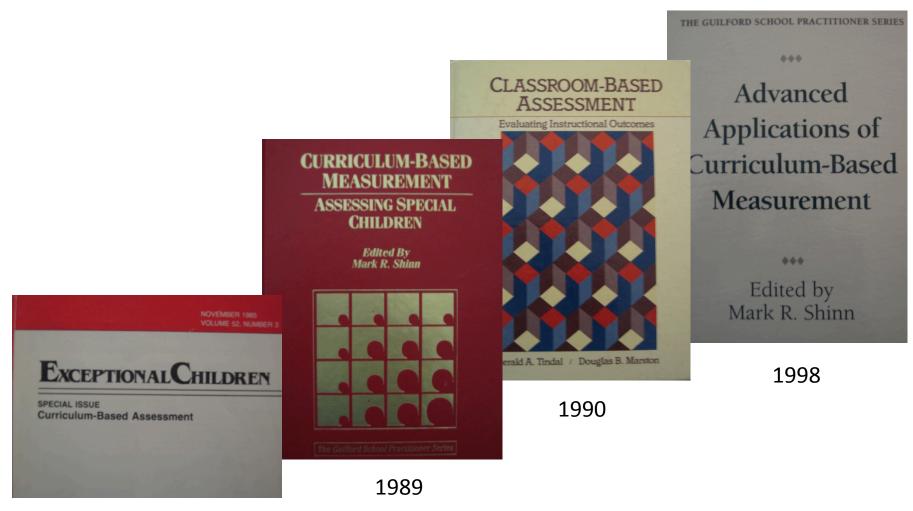


And then there was research....





And eventually proliferation...





Various definitions appeared...

- Curriculum-Based Assessment Gickling (1985)
 27,800 references found in Google Scholar (119 in DBs)
- Curriculum-Based Measurement Deno (1985) 13,400 references found in Google Scholar (106 in DBs)
- Classroom-Based Assessment Tindal (1990) 30,700 references found in Google Scholar (120 in DBs)
- Curriculum-Based Evaluation— Howell (1999)
 27,000 references found in Google Scholar (77 in DBs)



Ironically, perspectives were ignored...

- Measurement principles de-emphasized (no tenets from National Council on Measurement in Education)
- Scaling never addressed (though alternate forms considered) and therefore no items or attention to IRT
- Normative performance implicit (and without sampling plans) and lack attention to standard setting technology
- Time series graphs presented but no other graphic displays
- Integrated decision making implicit (at systems above classroom)



Note to Field on Validity

- Messick Evidence and consequence
- Standards for Educational and Psychological Testing — Propositions and Interpretations
- Gersten Validity as instructional practice



Messick ala 1994

- "The construct basis of test validity includes evidence and rationales for evaluating the intended and unintended consequences of test interpretation and use in both the short and the long term" (p. 21).
- Particularly prominent is the evaluation of any adverse consequences for individuals and groups that are associated with bias in test scoring and interpretation or with unfairness in test use" (p. 21).

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13-23.



Messick ala 1995

- "Validity is an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on test scores or other modes of assessment.
- Validity is not a property of the test or assessment as such, but rather of the meaning of the test scores...and a function of the persons responding as well as the context of the assessment.



Messick, S. (1995). Standards of validity and the validity of standards in performance assessments. *Educational Measurement: Issues and Practice. Winter*, 5—8.

Educational Standards ala 1999

- "Validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by the proposed uses of tests."
- "The proposed interpretation refers to the construct or concepts the test is intended to measure."
- Clarified by propositions that support proposed interpretations, each of which may require different types of evidence.



Gersten ala 1995

 "Fully determining the validity of an assessment process transcends what any one researcher can accomplish. It is a task for a community of researchers and practitioners to consider meanings and utility of assessment procedures in relation to current thinking about how to improve instructional practice and issues raised by studies of implementation" (p. 512). Gersten, R., Keating, T., & Irvin, L. K.



(1995). The burden of proof: Validity as improvement of instructional practice. *Exceptional Children*, *61*(6), 510-519.

NC – RTI Evaluation Criteria

Validity of the Performance Level Score is the extent to which the score (or average/median of 2-3 scores) represents the underlying construct.

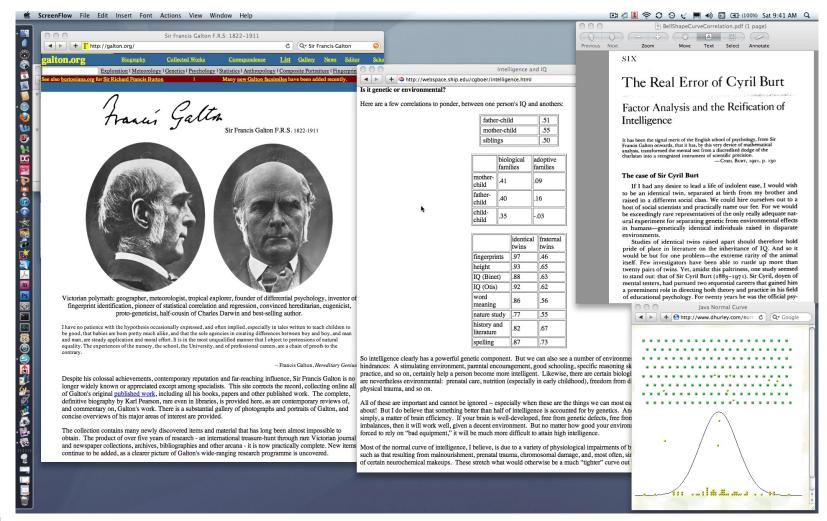
Rating Rating defined

- Convincing evidence: The validity for the performance level score (e.g., content, concurrent, predictive and/or construct) is adequate.
- *Unconvincing evidence*: The validity for the performance level score (e.g., content, concurrent, predictive and/or construct) is not adequate.
- *No evidence*: Validity for the performance level score data were not provided.

Rates of Improvement Specified as the slopes of improvement or average weekly increases, based on a line of best fit through the student's scores.



The Wisdom of Crowds





easyCBM Learning Management Systems

- Adaptability to fit district context with settings to control fields, resource allocation, and access
- Historical record of academic performance, progress, and intervention information
- Capacity to share student data seamlessly within data teams using differential levels of access
- Designed to facilitate sharing information and guide decision making for key stakeholders: teachers, administrators, specialists, parents, and students



easyCBM Grants

- Reliability and Validity Evidence for Progress Measures in Reading. U.S. Department of Education, Institute for Educational Sciences, Budget \$1,596,638 from June 2010 June 2014.
- Developing Middle School Mathematics Progress Monitoring Measures. U.S. Department of Education, Institute for Educational Sciences, Budget \$1,631,401 from June 2010 June 2014.
- Response to Intervention with Reading Curriculum-Based Measures. U.S.
 Office of Special Education Programs, Steppingstones of Technology
 Innovation for Children with Disabilities. Budget \$396,736 from May 2009
 April 2011.
- Assessments Aligned with Grade Level Content Standards and Scaled to Reflect Growth for Students with Disabilities (SWD) and Persistent Learning Problems (PLP). U.S. Department of Education, Institute for Educational Sciences. Budget \$1,525,552 from May 2007 – April 2011.
- Model Demonstration Centers on Progress Monitoring (CFDA 84.326M).
 U.S. Department of Education. Budget: \$1,189,790 from January 2006 December 2008.



easyCBM

- Free teacher version easy CBM
- Riverside district version easy CBM
- http://easyCBM.com/
- Disclosure on income to BRT for research and development with no personal remuneration
- Current use patterns
- Funded since 2006 with over \$6,000,000.





Three References for Personalizing Learning

- Norm Referenced Evaluation [NRE] How do students compare to others? (used for allocation of resources)
- Criterion Referenced Evaluation [CRE] How well do students perform? (used for diagnostics and targeting the content of instruction)
- Individual Referenced Evaluation [IRE] How much are students improving? (used for evaluating instructional programs – a.k.a. RTI)





Home

Students

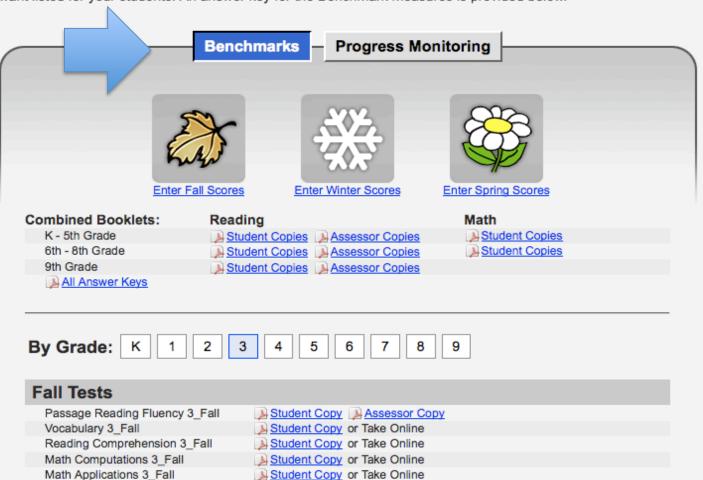
Measures

Reports

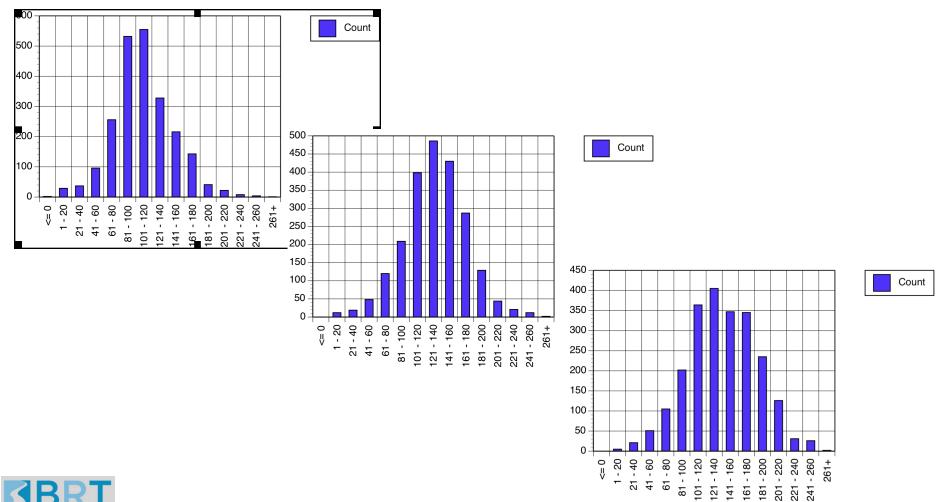
Admin Account

Measures on easyCBM

Select the grade level difficulty of the measure you want to access, then scroll to its section. For measures that can be administered online, have your students go to http://4j.or.easycbm.com/brtadmin and follow the on-screen instructions. For Benchmark measures, your district admin must enable them before they'll show up for your students. For Progress Monitoring measures, mark the checkbox next to each measure you want listed for your students. An answer key for the Benchmark measures is provided below.



Norm-Referenced Evaluation [Oral Reading Fluency]





Defining Risk and Allocation of Resources

Individual Measures	Risk Ratings
= 90 - 100 Percentile = 21 - 89 Percentile = 11 - 20 Percentile = 0 - 10 Percentile	Three Tests Risk Two Tests Risk One Test Risk Low Low Some Some Some Some High Some High Some High High High High High



Account



Home

Students

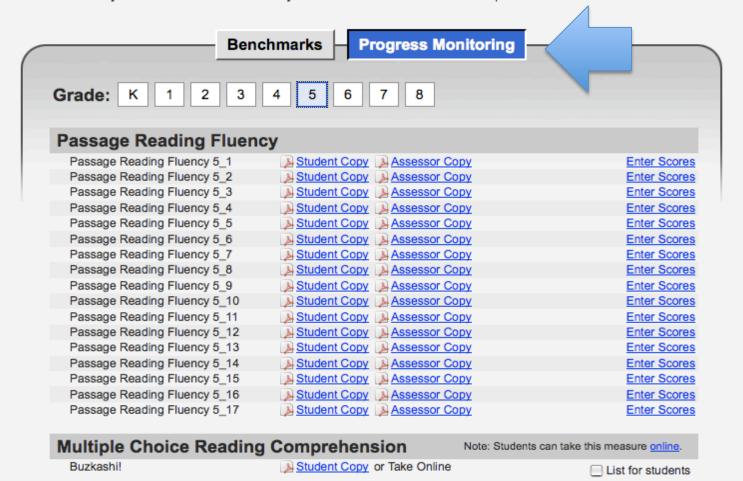
Measures

Reports

Admin

Measures on easyCBM

Select the grade level difficulty of the measure you want to access, then scroll to its section. For measures that can be administered online, have your students go to http://4j.or.easycbm.com/brtadmin and follow the on-screen instructions. For Benchmark measures, your district admin must enable them before they'll show up for your students. For Progress Monitoring measures, mark the checkbox next to each measure you want listed for your students. An answer key for the Benchmark measures is provided below.





Home

Students

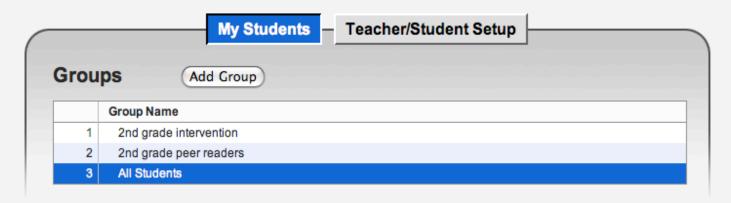
Measures

Reports

Account

Managing your Students

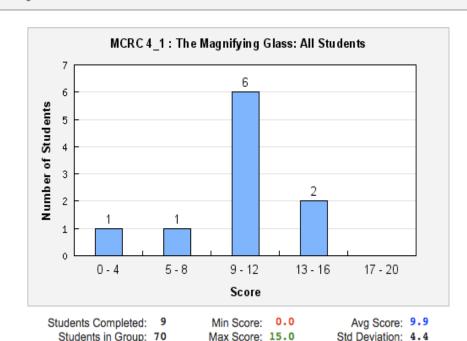
Add new Students or Groups by clicking the Add Student or Add Group buttons. If a student is in the currently selected group, they will have a check next to their name. If the student is not in the group, their checkbox will be unchecked. When a grouping or student is selected, links will appear that allow you to edit the entry.



Students (Create Temporary	Record

	In Group	Student Name
1	▼	Adalberto Ball
2	\checkmark	Alaina Bernier
3	\checkmark	Bobbie Dimauro
4	\checkmark	Darline Engstrom
5	\checkmark	Perry Leiser
6	\checkmark	Rusty Macy
7	\checkmark	Reatha Nelson
8	\checkmark	Ross Pitcher
9	$ \overline{\checkmark} $	Christopher Poore

Summary



Item Analysis

Top Easiest									
Item	Type	Type Students Correct							
6	Inferential	8 of 9	89%						
8	Literal	7 of 9	78%						
10	Evaluative	7 of 9	78%						
13	Evaluative	6 of 9	67%						
1	Literal	6 of 9	67%						
11	Literal	6 of 9	67%						
17	Literal	6 of 9	67%						
7	Evaluative	6 of 9	67%						
4	Inferential	6 of 9	67%						
5	Literal	6 of 9	67%						

Top Hardest								
ltem	Type	Type Students Correct						
14	Literal	2 of 9	22%					
16	Evaluative	2 of 9	22%					
9	Inferential	3 of 9	33%					
19	Inferential	3 of 9	33%					
15	Inferential	3 of 9	33%					
2	Literal	4 of 9	44%					
20	Evaluative	4 of 9	44%					
3	Inferential	4 of 9	44%					
12	Inferential	5 of 9	56%					
18	Evaluative	5 of 9	56%					

Group Report:

provides information helpful for grouping students and insight into the item types on which they need more work

Item-Person Maps

Item Analysis with Student Names

All item analysis reports have been updated to show a single table of items from easiest to hardest, and they now have a column for the names of the students who got the corresponding item incorrect. Hovering your mouse cursor over a name will highlight the same name for other items as well.

Item Analysis

Easiest to Hardest Items									
ltem	Type	Students Correct	Percentage	Student Names, Incorrect					
6	Inferential	4 of 4	100%						
5	Literal	4 of 4	100%						
1	Literal	4 of 4	100%						
12	Inferential	3 of 4	75%	Christia T					
19	Literal	3 of 4	75%	Denver S					
14	Literal	3 of 4	75%	Denyse F					
16	Evaluative	3 of 4	75%	Christia T					
15	Inferential	3 of 4	75%	Denyse F					
8	Literal	3 of 4	75%	Denyse F					
3	Inferential	3 of 4	75%	Christia T					
2	Literal	3 of 4	75%	Christia T					
17	Inferential	3 of 4	75%	Christia T					
4	Inferential	2 of 4	50%	Gail P, Christia T					
7	Evaluative	2 of 4	50%	Denyse F, Christia T					
20	Evaluative	2 of 4	50%	Denyse F, Denver S					
18	Evaluative	1 of 4	25%	Gail P, Denver S, Christia T					
10	Evaluative	1 of 4	25%	Gail P, Denver S, Christia T					
9	Inferential	1 of 4	25%	Denyse F, Gail P, Christia T					
11	Literal	0 of 4	0%	Denyse F, Gail P, Denver S, Chris					
13	Evaluative	0 of 4	0%	Denyse F, Gail P, Denver S, Chris					



Interventions



Logged in: tealy (logout)

1/19/2010 - 1:38:52 pm

Home

Students

Measures

Reports

Account

Reports and Analysis

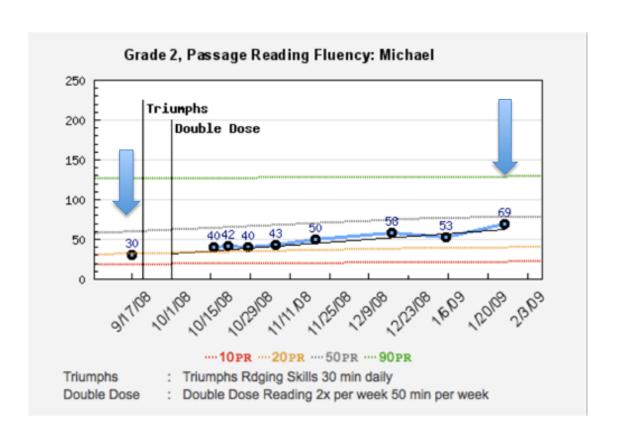


To view a **Group** report, click on the name of the group, and all of their active CBMs will appear below. Select a CBM name to see a summary and list of student scores. Then click "View" to see any student's actual submission. Select the **Individuals** subsection for easy one-click access to system wide data by student. Interventions are now accessible under the Individuals subsection, on the right-hand side of the table.

			ons for Ruthanne Almendarez	
Date	Subject	Label	Description	
10/16/2008	Reading	Tier 3	Instructor: Certified Teacher Curriculum:Phonics/PA Number of sessions/week:1 Length of session:20 minutes Size of group:5 PM measure:LN/LS Frequency of PM:every 2 weeks	Edit Delete
11/15/2008	Reading	Change	Increase to 2 days per week.	Edit Delete
1/30/2009	Reading	Change	Increase to five days per week with certified teacher	Edit Delete
3/11/2009	Reading	Change	Decrease to two days per week with certified teacher and focus on PA and segmentation	Edit Delete
3/19/2009	Reading	Change	Increase to three days a week with IA	Edit Delete

New Intervention

Individual Student Report

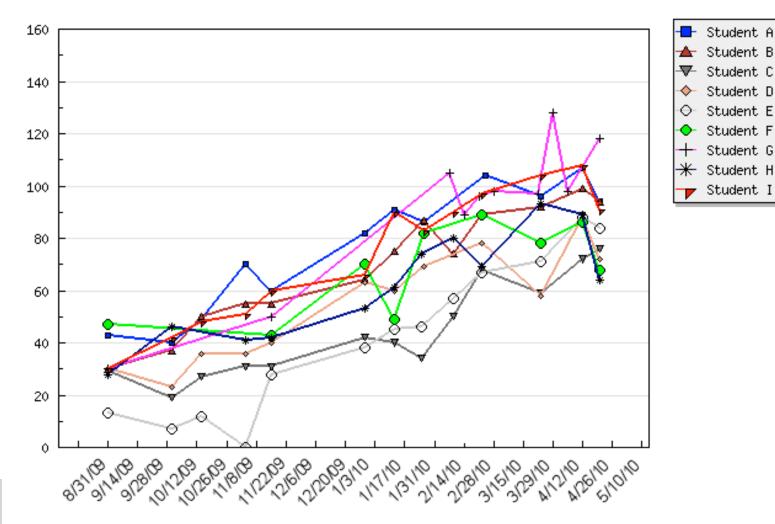


This report provides information helpful for judging the effectiveness of interventions for a particular student.



Response to Intervention

Group Passage Reading Fluency Performance (Only shown for groups of 10 students or less)





Summative Risk Analysis

Risk Analysis Report

This report compares student risk ratings for fall vs winter, winter vs spring, and fall vs spring. Here is an example of the report when viewing a particular grade, building-wide:

Grade K Risk Analysis

Risk Level 🔺
Risk 0
Risk 1
Risk 2
Risk 3
Risk 4
Risk 5
Risk 6
Totals

Fall	Winter	Change
49%	47%	2%↓
6%	29%	23%↑
18%	9%	9%↓
6%	9%	3%↑
12%	3%	9%↓
3%	1%	2%↓
7%	1%	6%↓
100%	100%	-

Winter	Spring	Change
47%	97%	50%↑
29%	3%	26%↓
9%	0%	9%↓
9%	0%	9%↓
3%	0%	3%↓
1%	0%	1%↓
1%	0%	1%↓
100%	100%	-

Fall	Spring	Change
49%	97%	48%↑
6%	3%	3%↓
18%	0%	18%↓
6%	0%	6%↓
12%	0%	12%↓
3%	0%	3%↓
7%	0%	7%↓
100%	100%	-

Options: Total | Percentage

Intact | Cohort

And here is the report when viewing a teacher's class list:

Grade 3 Risk Analysis

	Student Name	Fall	Winter	Change		Winter	Spring	Change	Fall	Spring	Change
1	Boulton, Fanny	0	0	-		0	-	-	0	-	-
2	Canterbury, Earle	0	0	-		0	-	-	0	-	-
3	Catlett, Michal	2	2	-		2	-	-	2	-	-
4	Colley, Milford	6	2	4↓		2	-	-	6	-	-
5	Fambro, Cristin	-	0	-		0	-	-	-	-	-
6	Fawcett, Lenard	0	0	-		0	-	-	0	-	-
7	Gamache, Joane	3	0	3↓		0	-	-	3	-	-
8	Kan, Virgil	1	2	1↑		2	-	-	1	-	-
			_	4.	1	_					



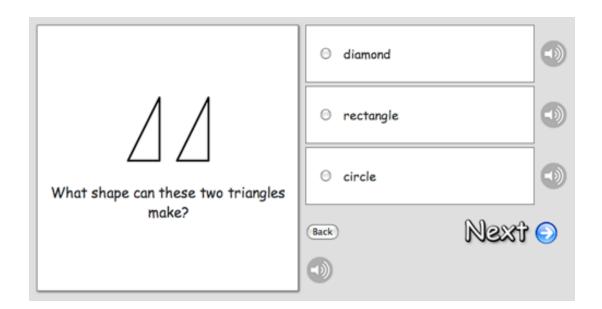
New Vocabulary Measures

Last year we only had two vocabulary measures per grade, one for fall and one for spring. Now we have brand new vocabulary measures for fall, winter, and spring, as well as 10 progress monitoring measures per grade, for grades 2 - 8! Since the measures are so new, we won't have percentiles or color-codings for vocab until late October for fall, late February for winter, and late June for spring. The vocab measure won't be included in risk calculations until then as well.

- 7. Harry filled the goblet with drinking water. What does goblet mean?
 - A. bowl B. plate C. glass
- 8. The digits on your hand are your:
 - A. numbers B. fingers C. toes
- 9. chocolate is a:
 - A. plant B. diet C. flavor

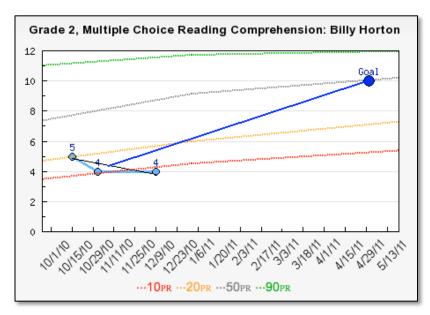
More Read Aloud for Math

Shortly after we added the online math measures to the system, we created a read aloud option for the question part of the kindergarten and first grade tests. This year we've added read aloud options for all questions and all answer options that have text, for all grades K through 8, and for both progress monitoring and benchmarks.



Goal Lines for Individual Graphs

You now have the ability to set goal data points for your students, one for each measure type at a grade level. It will show up as a big dark blue dot on the graph, at the date and score that you set. Once the student has taken at least three measures, a line will be drawn from the midpoint of those scores to the goal you've set. We hope this feature will help teachers quickly determine if students are making adequate progress or not.



Customized Measures of Learning

- Dynamic Measures
 - Mathematics: Sequencing, Simple Operations
 - Reading Word Lists
- Teacher Made Measures
 - Oral Reading Fluency
 - Maze Vocabulary
- http://www.brtprojects.org/labs



On Two Metaphors for Learning and the Dangers of Choosing Just One

"Educational research is caught between two
metaphors...the acquisition metaphor and the
participation metaphor. Both of these metaphors are
simultaneously present in most recent texts, but
while the acquisition metaphor is likely to be more
prominent in older writings, more recent studies are
often dominated by the participation metaphor" (p.
4).



Sfard, A.(1998). On Two Metaphors for Learning and the Dangers of choosing just one. *Educational Researcher*, 27(2), pp. 4—13.

Acquisition and participation become coins of the realm...

- Measurement of growth and attention to individual differences and making an individual difference (ala learning from an acquisition metaphor)
- Importance of human capital, structures and functions (ala learning from a participation metaphor)



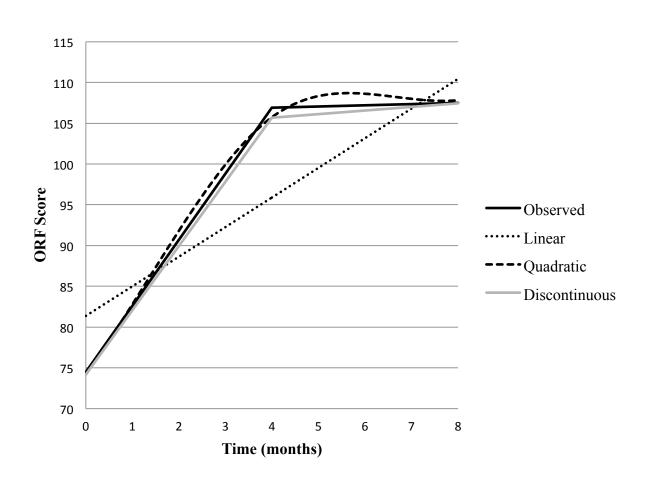
ORF Means and SDs by Grade*

	Fall	Winter	Spring
Grade	Mean (SD)	Mean (SD)	Mean (SD)
3	74.48 (35.41)	106.90 (39.94)	107.48 (39.24)
4	102.31 (36.58)	122.89 (38.43)	130.91 (41.49)
5	134.48 (42.81)	143.38 (39.72)	156.01 (40.18)

^{*}Nese, J. F. T., Biancarosa, G., Anderson, D., Lai, C. F., Alonzo, J., & Tindal, G. (in press). Within-year oral reading fluency with CBM: A comparison of models. *Reading and Writing: An Interdisciplinary Journal*.

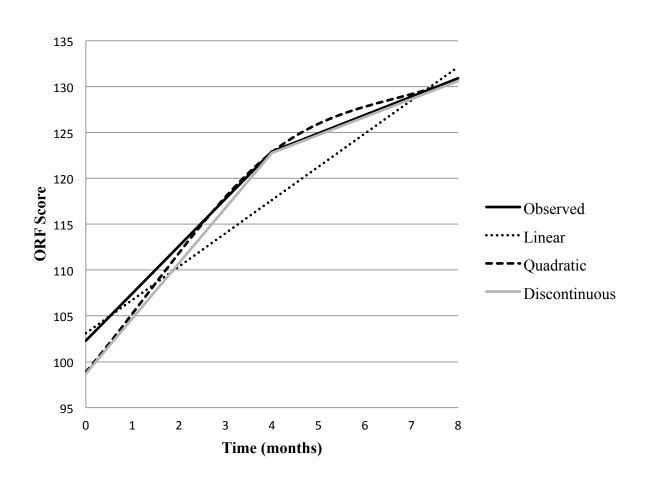


Predicted Trajectories for Three Parameterizations of Growth Compared to Observed Growth in Grade 3



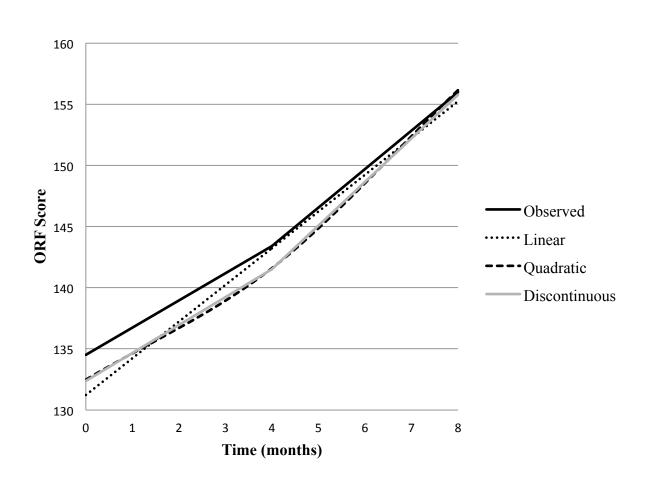


Predicted Trajectories for Three Parameterizations of Growth Compared to Observed Growth in Grade 4





Predicted Trajectories for Three Parameterizations of Growth Compared to Observed Growth in Grade 5





Learning in different student groups...

- Across grade levels, females began the year reading about 4 more cwpm than males.
- Students eligible for FRPL began the year reading about **11** cwpm less than students not eligible.
- SWD began the year reading about **37** cwpm less than general education students.
- LEP students began the year reading about 20 cwpm less than non-LEP students



Implications for students...

- "We also see variability in student easyCBM scores that seem to be related to their home life, lack of sleep or food, the time of day of testing or their emotional state on the days they test."
- "When there hasn't been enough progress over 12-18
 weeks of tier 3, or there have been enough data points
 below the 20th percentile, we look to see if it is
 possible to refer that student to special education."
- "There are still very low performing students; there is still flat achievement for some."



Implications for teachers...

- Adequate funding and highly skilled personnel seem to be our two biggest challenges
- "We have had little professional development relating to teaching reading skills."
- "I haven't received any training in several years about progress monitoring, or decision making either."
- As we can't find time to take students out of other core subjects like writing or math, Tier 3 must be done out of our reading block.



The nexus of acquisition and participation...

- In the focus groups, teachers referred to the usefulness of having access to this common data, regardless of school site.
- "easyCBM assessment system enables everyone in the district to use a common, streamlined benchmarking and progress monitoring assessment system that makes it easy to track interventions and their impact over time."



Classroom instruction gets arranged...

- We have a 90 minute reading block where all students receive tier one (adopted main curriculum) and some form of tier two (differentiated instruction).
- Students get in this time whole group, independent centers/work and small group instruction.
- Students under the 20th PR and Tier 3 get twice a week for 25-30 minutes outside the reading block.
- Split the kids up into skills groups that vary with the most needy group having far fewer students; groups are fluid
- We utilize three classroom teachers, 3 instructional assistants, and a retired volunteer teacher.
- Interventions are based on skills to develop within the group.



Student performance and progress data get used...

- PM every 2 weeks for students in the lowest 20th percentile.
- Whole class progress monitoring every 6 weeks.
- Grade level standards 3 times in a row pulled from Tier
- Data meetings are held at each grade level every 6 weeks.
- Rotating schedule for particular grade level to meet
- The structure: review the past meeting notes, discuss current data, and decide which students are to be given Tier 3 and how the Tier 2 needs to be adjusted.
- Make instructional decisions: changing curricula, changing groups, entering or exiting kids from Tier 3 services and RFA's (Requests for Assistance).
- Look first at students from the last meeting and then new students of concern that surface because of easyCBM data.



New roles get established...

- Data Scene Investigator (DSI)
- Student Achievement Coordinator (SAC)
- Title I teacher
- Instructional Intervention Progress Monitoring IIPM Response to Intervention (RTI) team members
- Learning Center (SPED) time
- TLC
- School psychologists



In the end...results from three centers

- MPLS: Key thing is the OCR website in the MSP as helpful and supportive
- PA: Changes in language and professional relationships (even though the original crew is no longer present; this is what you do, not a special project).
- PA and OR: Due to the _____ who maintained the general outline



Explanations and attributions...

- Data team and administrative teams with professional development; all said the principal was key factor and has added behavior support to building wide perspective.
- Teachers said that once they saw kids achieving and that kids were reading, it affirmed their value as a teacher; they see kids learn who they had given up on before.



Eventual attention to parts and whole...

- Principal takes perspective of ensuring teachers' needs (and interests) to support RTI: Advocate and someone to push for it and with power and influence.
- Though school leadership may not be the only critical level, as teachers are needed. Teachers appear to be more stable than leadership.



Other factors for RTI success...

- Schedule time and meetings: Make the trains run on time
- Delegate someone to run the meeting
- Provide strong school level leadership
- Secure extra supports
- Continue to use local resources



The end...

http://www.brtprojects.org/

