Evidentiary Requirements for Progress Monitoring Measures When Used for Response to Intervention

Roland H. Good III
University of Oregon
Dynamic Measurement Group

DIBELS Summit
Albuquerque, New Mexico
February 18, 2009

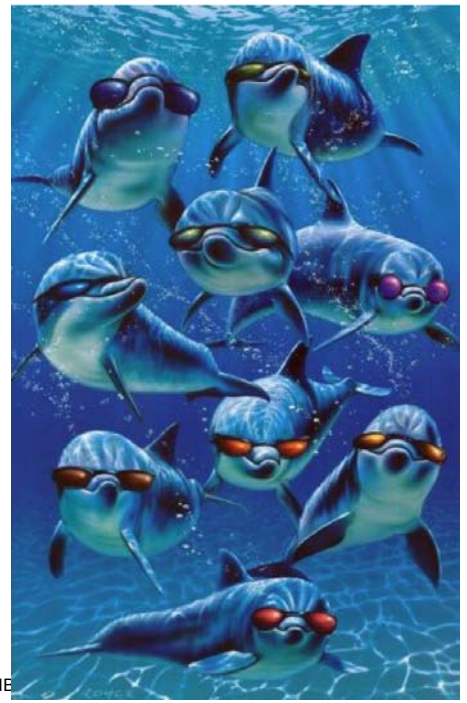
Acknowledgements

- Work in this presentation was conducted with the support of the University of Oregon, Center on Teaching and Learning, DIBELS Data System.
- The work and thinking represented in this presentation benefitted significantly from the comments, suggestions, and involvement of Ed Kame'enui, Scott Baker, Hank Fien, Mike Stoolmiller, and Kelli Cummings, with assistance from Courtney Wheeler, Rebecca Briggs, and Darci Burns.
- The conceptualizations of evidentiary concerns benefited significantly from the input of former and current members of the DIBELS/IDEL Research Team at the University of Oregon.
- My colleagues at Dynamic Measurement Group continue to provide the thinking and challenging perspectives that are invaluable to building this understanding of DIBELS.
- Opinions, conclusions, and recommendations are solely those of Roland Good unless otherwise indicated.

Data Source for Analyses and Normative Context

- University of Oregon DIBELS Data System, 2006-2007 academic year.
  - Focus on second grade students, first semester, beginning of year (BOY) to middle of year (MOY).

There are lots of cool porpoises



What is your purpose for implementing RTI?

- **Maximizing student learning?**
  - Prevention
  - Early intervention
  - Screening
  - Improving school-wide system of instruction
- **Eligibility?**
  - Special education
  - Learning disability
  - Labeling and classification
- **Both?**

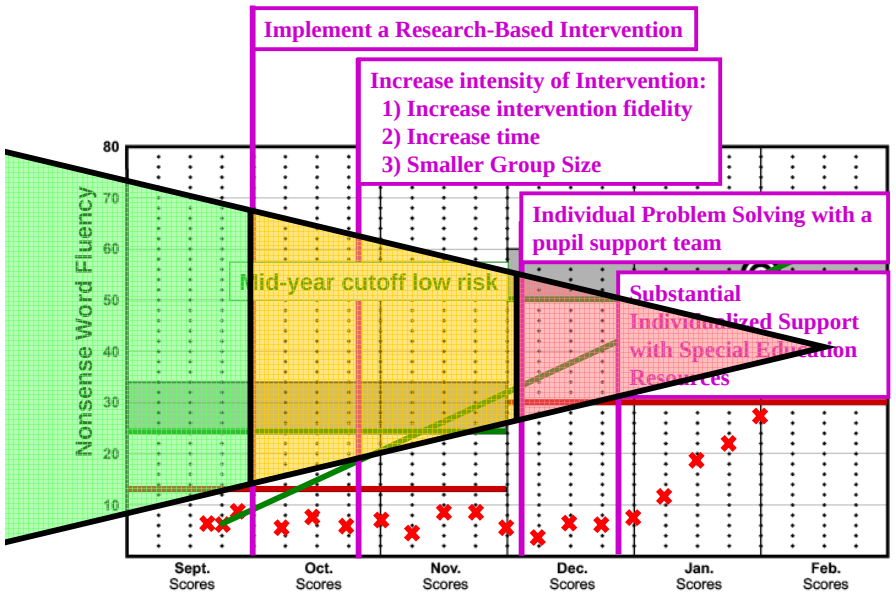
Implementing an RTI Model Only for Eligibility is not Enough

- Using RTI only for eligibility decisions is like buying a brand new Ferrari and never taking it out of first gear.
- Referral for special education eligibility evaluation because of academic difficulty is not an appropriate starting point for RTI.
- Eligibility based on lack of adequate progress is NOT a desirable endpoint for RTI.
- Response to intervention (RTI) in a prevention-oriented system of generally effective instruction (e.g., a three-tier model) IS a power approach to maximizing student learning and can provide a defensible basis for special education eligibility decisions when rigorous standards for the technical adequacy of the decisions are met.

Eligibility is a High Stakes Decision

- High Stakes Decisions - Eligibility
  - One-time decision point that is not easily modified.
  - Immediate life impact is likely.
    - Positive consequences – support, intervention.
    - Unintended negative consequences are likely – more restrictive environment, stigmatization.
  - High stakes decisions require a higher degree of rigor in evidentiary considerations.
- Low Stakes Decisions – Maximizing learning
  - Set of ongoing decisions
  - Self-correcting decisions. Initial decisions are monitored and re-evaluated on an ongoing basis with adjustments as necessary.
  - Gradual life impact is likely.
    - Gradual onset of positive consequences
    - Minimize unintended negative consequences
  - Low stakes decisions may be made with a lower degree of rigor in evidentiary considerations.

Outcomes Driven Model and RTI



Three Crucial Measurement Decisions in RTI

1. Is the student making adequate year-to-year progress?
   - Maximize learning: Is the student making adequate progress toward meaningful long term goals?
   - Eligibility: Does the student have severe low achievement that may indicate learning difficulty?
2. Is the student receiving generally effective instruction?
3. Is the student making adequate week-to-week progress?
   - Maximize Learning: Is the student making adequate progress?
   - Eligibility: Does the student display a serious, stubborn, sustained lack of adequate progress when provided with generally effective instruction?

Reliability Evidence Required for Defensible Educational Decisions

- Reliability – Decisions should be reasonably stable across trivial changes in conditions.
- Thou shalt not make capricious decisions about children.
  - Maximize Learning: lower standard because decisions are self-correcting and low stakes.
  - Eligibility: Rigorous standards because high stakes decisions.
    - Decisions about Level: reliability of .90 or higher.
    - Decisions about Rate of Progress: No specific standards or criteria are generally accepted. More reliable is important.

Normative Context Required for Defensible Educational Decisions

- Normative context – How well is the student performing compared to a relevant comparison group.
  - Local norms compare performance other children in the student's classroom, school, or district.
  - National norms compare performance to other children around the nation.
  - Other specific comparison groups.
- Maximizing Learning: What are reasonable expectations for grade level peers?
- Eligibility: If almost everyone has it, doesn't have it, does it, or can't do it, then it is not a disability and not evidence for eligibility for special education.

Defensible Educational Decisions Require Evidence the Skills are Meaningful

- Meaningful evidence links decisions to outcomes. Reschly would call this the Outcomes Criterion.
  - Prognosis: Students with a particular level of skills or educational needs have lower likelihood of favorable outcomes.
  - Dosage: Students with lower likelihood of favorable outcomes benefit from more instructional time.
  - Intervention: When students with a particular set of skills or educational needs are provided with a specific intervention their outcomes are better than if they receive a different intervention.
- How important is the difference in outcomes. Would a parent care?

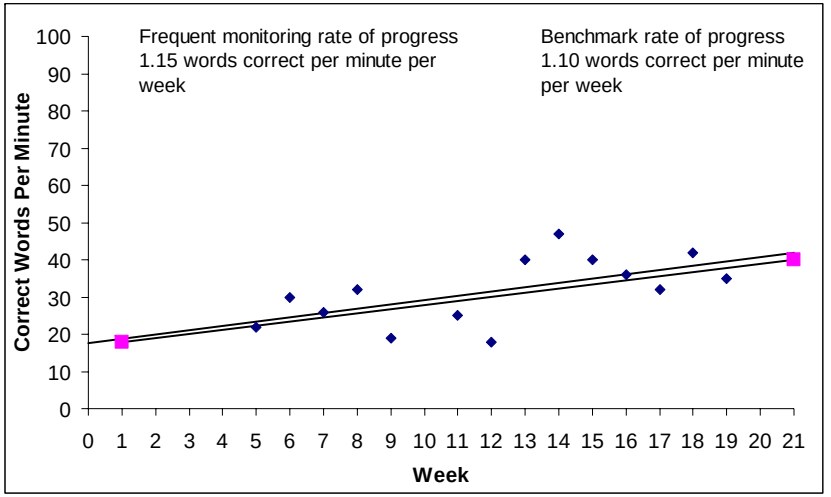
Evidentiary Requirements for RTI

Evidentiary Considerations for the Educational Decisions Required for Response to Intervention Models

Educational Decision	Evidentiary Consideration: Reliable	Normative Context	Meaningful
1. Is the student making adequate year-to-year progress?	X	X	X
2. Is the student receiving generally effective instruction?	?	?	?
3. Is the student making adequate week-to-week progress?	+/-	+/-	+/-

Note. X = generally strong and persuasive evidence. ? = level of evidence is unestablished. +/- = emerging evidence base.

Benchmark Rate of Progress and Frequent Monitoring Rate of Progress



Descriptive Statistics for and Correlations Between Benchmark Rates of Progress and Frequent Monitoring Rates of Progress by Number of Progress Monitoring Data Points for Students from Beginning of Year to Middle of Year, Second Grade, 2006-2007

Exact number of PM points	Number of students	Benchmark rate of progress *M*	Benchmark rate of progress *SD*	Frequent monitoring rate of progress *M*	Frequent monitoring rate of progress *SD*	Reliability of frequent monitoring rate of progress	Correlation Benchmark progress with Frequent monitoring progress
8 points	4529	1.39	0.75	1.07	1.11	.31	.45
10 points	2251	1.39	0.73	1.02	0.96	.34	.37
12 points	1278	1.36	0.75	1.07	0.86	.36	.48
14 points	471	1.39	0.72	1.39	0.93	.60	.54
16 points	82	1.45	0.70	1.36	0.85	.61	.56
18 points	27	1.12	0.69	1.41	0.77	.68	.60

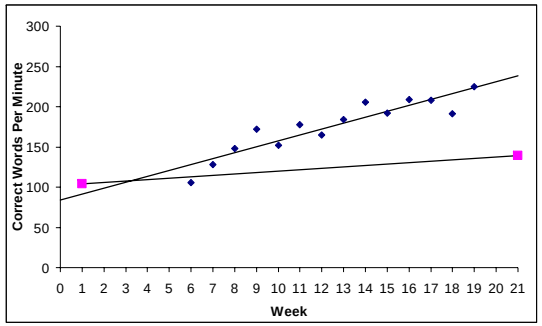
Reliability of Rates of Progress Combined Progress Monitoring and Benchmark

- Reliability of Rate of Progress for
  - 886 students with 14 or more progress monitoring assessments plus 2 benchmark assessments

.64

Even combining Benchmark and Progress Monitoring, the reliability of rate of progress does not fill us with confidence. What's going on?

Some Goofy Patterns



- Which is more accurate for this student: Benchmark Rate of Progress or Frequent Monitoring Rate of Progress?
- Why are we monitoring this second grade student's progress anyway?

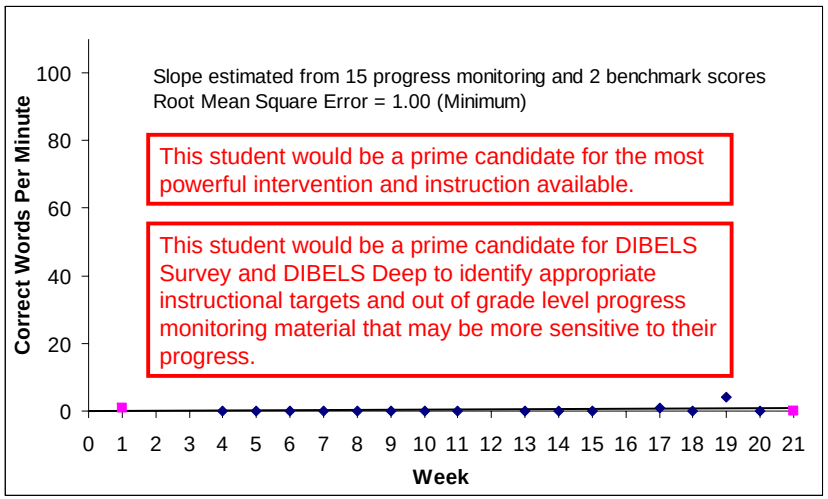
Reliability of Rates of Progress Combined Progress Monitoring and Benchmark

- Reliability of Rate of Progress for
  - 886 students with 14 or more progress monitoring assessments plus 2 benchmark assessments .64
  - 745 students with 14 or more progress monitoring assessments plus 2 benchmark assessments who were Intensive or Strategic at beginning of year. .67

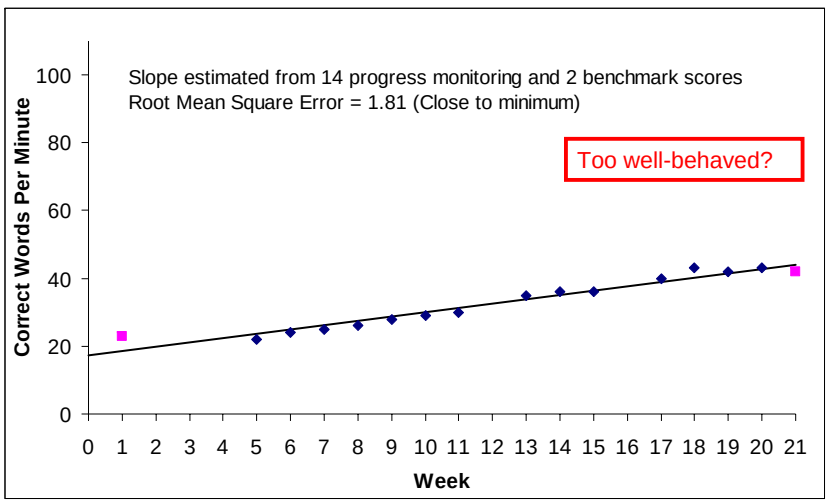
Even combining Benchmark and Progress Monitoring and only including students who were in the Intensive or Strategic range at the beginning of year second grade, the reliability of rate of progress does not fill us with confidence. Is something else going on?

Maybe some progress monitoring is not "well behaved".

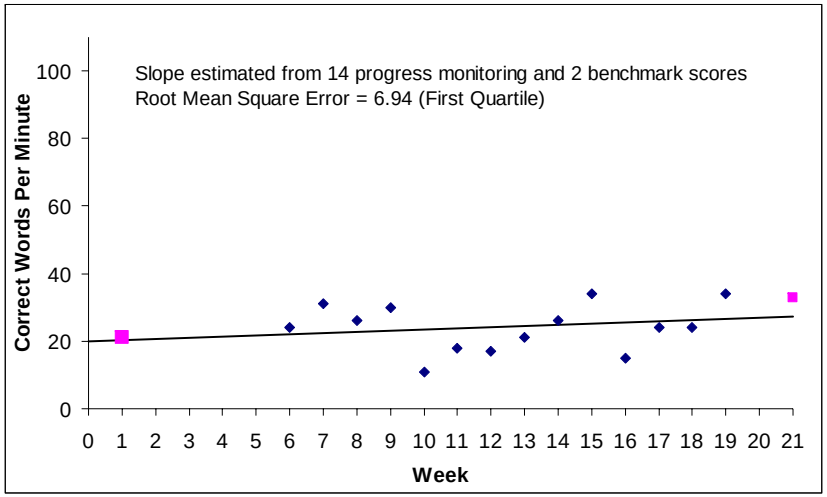
Minimum Bounce
Extremely Well-Behaved Progress Monitoring



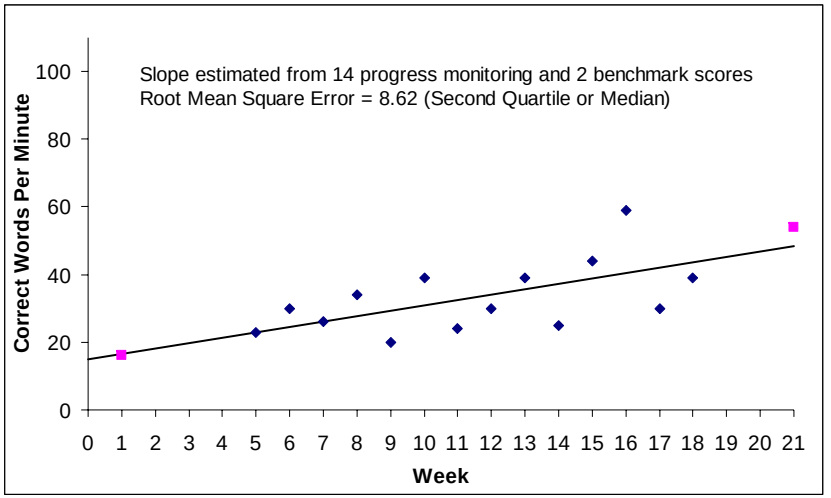
Close to Minimum Bounce
Extremely Well-Behaved Progress Monitoring



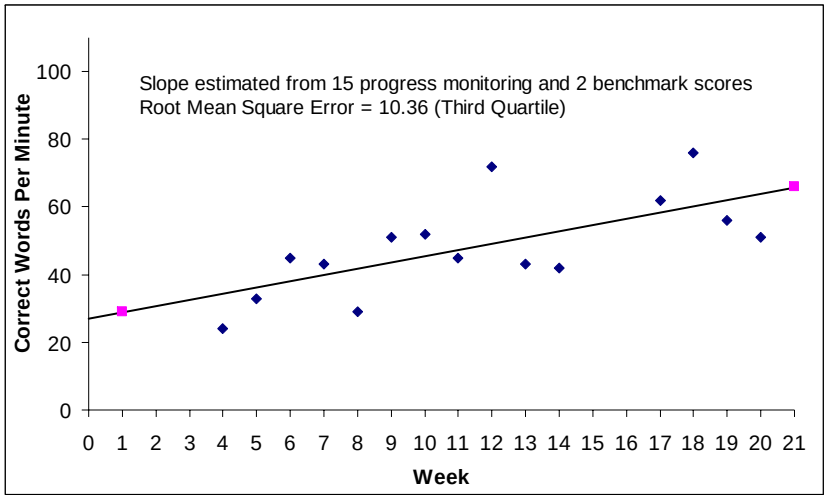
First Quartile (25th Percentile) in Bounce
Very Well-Behaved Progress Monitoring



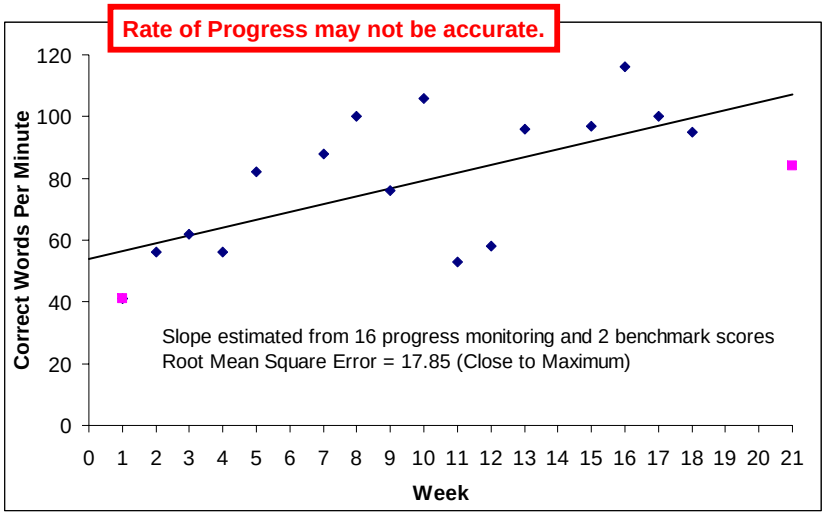
Median Bounce
Typically Behaved Progress Monitoring



Third Quartile of Bounce
Upper Limit for Reasonable Rate of Progress?



Close to Maximum Bounce
Poorly Behaved Progress Monitoring



Reliability of Rates of Progress Combined Progress Monitoring and Benchmark

- Reliability of Rate of Progress for
  - 886 students with 14 or more progress monitoring assessments plus 2 benchmark assessments — .64
  - 745 students with 14 or more progress monitoring assessments plus 2 benchmark assessments who were Intensive or Strategic at beginning of year. — .67
  - 558 students with 14 or more progress monitoring assessments plus 2 benchmark assessments who were Intensive or Strategic at beginning of year and whose root mean square error (RMSE) was less than 10.36, the third quartile of rmse. — .75
  - 372 students with 14 or more progress monitoring assessments plus 2 benchmark assessments who were Intensive or Strategic at beginning of year and whose RMSE was less than 8.62, the second quartile of RMSE. — .78

Recommendations for More Reliable Rates of Progress for Eligibility Decisions

- Make sure assessments are accurate: Observe testers with *Assessment Integrity Checklists*.
- Use at least 14 progress monitoring assessments.
- Combined progress monitoring and benchmark assessments.
- Examine progress graph.
  - Is there high variability in frequent monitoring assessments? You may not have an accurate and reliable measure of rate of progress. There may be some other issue that is interfering with performance. Health? Conduct? Motivation?
  - If frequent monitoring assessments are reasonably well-behaved and provide coherent and consistent evidence of progress (RMSE < 10.36 on DORF) the rate of progress reliability will be about .75.
- The emphasis is on making good and defensible decisions.

Providing a Normative Context for Rates of Progress

- First, select a normative context for comparison:
  - *Typical rates of progress in a local normative context*, for example other students in the same classroom, school, or district.
  - *Typical rates of progress in a national normative context.*
  - *Expected rates of progress* for students in a national sample of schools with high rates of adequate progress.
  - *Possible rates of progress* for students receiving the highest quality interventions possible.

Typical in **Local Normative** Context

- Advantages
  - Examines student progress compared to other students with similar opportunities to learn.
  - Normative comparison is perfectly representative of local SES, diversity, and demographics because it is the local context.
  - Local normative comparisons are more readily available than national.
- Disadvantages
  - Adequate progress in one context (e.g., classroom) may not be adequate progress in another context (e.g., another classroom or school).
  - Local context may provide generally effective instruction, or may not.
  - May not be able to distinguish lack of progress due to a learning difficulty from lack of progress due to ineffective instruction.

Typical in **National Normative** Context

- Advantages
  - Provides an indication of the rates of progress achieved in typical curriculum and instruction.
  - Decisions about adequate progress (lack of adequate progress) are not context specific.
  - Able to examine typical rates of progress given different levels of initial skills.
  - Very large and therefore more stable normative context.
- Disadvantages
  - Difficult to obtain.
  - Typical curriculum and instruction in a national context may not be "generally effective" for students from diverse backgrounds or low skills.
  - Does not provide impetus for continuous, systems-level improvement at the school or district level.

Fourth Grade Reading Outcomes on the 2007 National Assessment of Educational Progress

Skill Level	Skill level definition	Nation (public) percent of fourth grade students *scoring below* (pp. 16, 52-53)	Nation (public) percent of fourth grade students from diverse backgrounds *scoring below* (pp. 54 & 57)
Basic	Basic denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at a given grade.	34%	54%, 51%, 49%, 50%
Proficient	Proficient represents solid academic performance. Students reaching this level have demonstrated competency over challenging subject matter.	68%	86%, 83%, 80%, 83%

Note: Students from diverse backgrounds includes students identified as Black, Hispanic, American Indian/Alaska Native, and eligible for free/reduced-price school lunch. From data reported in Lee, Grigg, & Donahue (2007).

Expected in a national sample of schools with high rates of adequate progress

- Advantages
  - Provides an indication of the rates of progress that could reasonably be expected in schools with high rates of adequate progress.
  - Attainable by schools: 1/3 of schools have high rates of adequate progress in each tier, about 10% of schools have high rates of adequate progress in all tiers.
- Disadvantages
  - Smaller normative context
  - Numbers of students in initial skill bands may be smaller.

Schools with High Rates of Adequate Progress for Intensive, Strategic, and Barely Benchmark Students

- Using school-based norms, a typical (median) school supports 14% of students who were intensive at BOY to progress to strategic or benchmark by MOY. Schools that support 23% are in the upper third.
- A typical (median) school supports 43% of students who were strategic at BOY to progress to benchmark by MOY. Schools that support 54% are in the upper third.
- A typical (median) school supports 93% of students who were barely benchmark (within 20 points of the benchmark goal) at BOY to continue adequate progress to benchmark by MOY. Schools that support 95% have high rates of adequate progress for benchmark students.
- 783 Schools had high rates of adequate progress for all three tiers.

Table 14

School Based Percentile Ranks for the Beginning of the Year to the Middle of the Year of 2nd Grade and Schools with More Than 40 Students

	Initial Status - BOY			Adequate Progress					Outcome - MOY		
School Percentile	Percent Intensive	Percent Strategic	Percent Benchmark	Intensive Adequate Progress	Intensive Exceptional Progress	Strategic Adequate Progress	Barely Benchmark Adequate Progress	Percent Adequate Progress	Percent Deficient	Percent Emerging	Percent Established
1	0	7	19	0	0	0	50	28	1	2	24
5	3	12	28	0	0	11	67	39	4	5	36
10	6	15	34	0	0	17	75	45	7	7	42
15	7	17	38	0	0	21	80	49	9	8	46
20	9	18	41	0	0	25	83	53	10	9	50
25	10	20	43	5	0	29	85	56	12	10	53
30	12	21	46	7	0	31	88	59	13	11	56
35	13	22	48	8	0	33	89	62	15	12	59
40	15	23	51	10	0	37	90	64	16	13	61
45	16	24	53	13	0	40	92	66	18	13	63
50	18	25	55	14	0	43	93	68	20	14	65
55	19	26	57	17	0	45	94	70	21	15	68
60	21	27	60	18	0	48	95	72	23	16	70
65	23	28	62	20	0	50	97	74	25	16	72
70	25	29	64	23	3	54	100	76	27	17	74
75	27	30	67	25	5	57	100	78	30	18	76
80	30	32	69	29	7	60	100	80	32	19	78
85	33	33	73	33	10	64	100	83	36	20	81
90	38	35	76	40	13	69	100	86	40	22	84
95	44	38	82	50	20	77	100	90	47	24	88
99	57	44	89	73	38	90	100	94	60	29	94

Note. Based on 6958 schools with 78176 students with beginning of year second grade ORF scores and middle of year second grade ORF scores.

Possible rates of progress for students receiving the high quality interventions

- Advantages
  - Provides an indication of what would be possible for a student to achieve with powerful intervention.
- Disadvantages
  - May require the training and support of research contexts.
  - May require that we hire Diana Prince as the interventionist (Wonder Woman).
  - May require resources beyond those generally available in schools.
  - May not be broadly sustainable.

Table 23

Zones of Growth by Level of Initial ORF Score in Beginning of Year Second Grade to Middle of the Year for Schools with 40 or More Students with High Rates of Adequate Progress for All Three Tiers (Conditional Probability of Intensive Reaching Strategic or Benchmark >= 23 and Conditional Probability of Strategic Reaching Benchmark >= 54 and Conditional Probability of Barely Benchmark Staying at Benchmark >= 95)

		BOY - MOY growth percentile			
BOY ORF	*n*	20th percentile	40th percentile	60th percentile	80th percentile
Intensive					
0 to 5	934	0.11	0.33	0.56	0.98
6 to 15	3145	0.40	0.70	1.05	1.53
16 to 25	6270	0.95	1.43	1.78	2.20
Strategic					
26 to 34	7862	1.30	1.73	2.06	2.43
35 to 43	7415	1.50	1.83	2.11	2.50

Can we Use Benchmark Rate of Progress to Provide a Normative Context for Frequent Progress Monitoring?

- We have a very large and representative sample of Benchmark assessments, the sample of frequent progress monitoring assessments is much smaller.
- With 14 frequent monitoring assessments, Benchmark rate of progress and frequent monitoring rate of progress correlate .54, just about the limits of the reliability of each estimate of rate of progress.
- Aggregating over many students can provide a more stable estimate of a mean rate of progress (or the 20th percentile of rate of progress).
- Do aggregates of frequent monitoring assessments correspond to aggregates of benchmark assessment?

Aggregating Rates of Progress in Blocks of Size 2

Student	Benchmark Rate of Progress	Frequent Monitoring Rate of Progress	Average Rate of Progress	Block	Average Benchmark Rate of Progress	Average Frequent Monitoring Rate of Progress
Student 1	0.20	0.17	0.18	Block 1	0.28	0.31
Student 2	0.35	0.45	0.40			
Student 3	0.75	0.45	0.60	Block 2	0.75	0.65
Student 4	0.75	0.85	0.80			
Student 5	1.35	0.66	1.00	Block 3	1.35	0.77
Student 6	1.35	0.89	1.12			
Student 7	1.45	1.05	1.25	Block 4	1.38	1.27
Student 8	1.30	1.49	1.40			
Student 9	1.25	1.94	1.60	Block 5	1.55	1.84
Student 10	1.85	1.74	1.80			
Student 11	2.10	1.89	2.00	Block 6	2.08	2.12
Student 12	2.05	2.35	2.20			
Student 13	2.30	2.50	2.40	Block 7	2.38	2.64
Student 14	2.45	2.78	2.62			
Student 15	2.55	3.10	2.83	Block 8	2.70	3.13
Student 16	2.85	3.16	3.00			

Aggregating Rates of Progress in Blocks of Size 3

Student	Benchmark Rate of Progress	Frequent Monitoring Rate of Progress	Average Rate of Progress	Block	Average Benchmark Rate of Progress	Average Frequent Monitoring Rate of Progress
Student 1	0.20	0.17	0.18			
Student 2	0.35	0.45	0.40	Block 1	0.43	0.36
Student 3	0.75	0.45	0.60			
Student 4	0.75	0.85	0.80			
Student 5	1.35	0.66	1.00	Block 2	1.15	0.80
Student 6	1.35	0.89	1.12			
Student 7	1.45	1.05	1.25			
Student 8	1.30	1.49	1.40	Block 3	1.33	1.49
Student 9	1.25	1.94	1.60			
Student 10	1.85	1.74	1.80			
Student 11	2.10	1.89	2.00	Block 4	2.00	1.99
Student 12	2.05	2.35	2.20			
Student 13	2.30	2.50	2.40			
Student 14	2.45	2.78	2.62	Block 5	2.43	2.80
Student 15	2.55	3.10	2.83			
Student 16	2.85	3.16	3.00			

Aggregating Rates of Progress in Blocks of Size 4

Student	Benchmark Rate of Progress	Frequent Monitoring Rate of Progress	Average Rate of Progress	Block	Average Benchmark Rate of Progress	Average Frequent Monitoring Rate of Progress
Student 1	0.20	0.17	0.18			
Student 2	0.35	0.45	0.40	Block 1	0.51	0.48
Student 3	0.75	0.45	0.60			
Student 4	0.75	0.85	0.80			
Student 5	1.35	0.66	1.00			
Student 6	1.35	0.89	1.12	Block 2	1.36	1.02
Student 7	1.45	1.05	1.25			
Student 8	1.30	1.49	1.40			
Student 9	1.25	1.94	1.60			
Student 10	1.85	1.74	1.80	Block 3	1.81	1.98
Student 11	2.10	1.89	2.00			
Student 12	2.05	2.35	2.20			
Student 13	2.30	2.50	2.40			
Student 14	2.45	2.78	2.62	Block 4	2.54	2.89
Student 15	2.55	3.10	2.83			
Student 16	2.85	3.16	3.00			

Aggregating Rates of Progress in Blocks of Size 5

Student	Benchmark Rate of Progress	Frequent Monitoring Rate of Progress	Average Rate of Progress	Block	Average Benchmark Rate of Progress	Average Frequent Monitoring Rate of Progress
Student 1	0.20	0.17	0.18			
Student 2	0.35	0.45	0.40			
Student 3	0.75	0.45	0.60	Block 1	0.68	0.52
Student 4	0.75	0.85	0.80			
Student 5	1.35	0.66	1.00			
Student 6	1.35	0.89	1.12			
Student 7	1.45	1.05	1.25			
Student 8	1.30	1.49	1.40	Block 2	1.44	1.42
Student 9	1.25	1.94	1.60			
Student 10	1.85	1.74	1.80			
Student 11	2.10	1.89	2.00			
Student 12	2.05	2.35	2.20			
Student 13	2.30	2.50	2.40	Block 3	2.29	2.53
Student 14	2.45	2.78	2.62			
Student 15	2.55	3.10	2.83			
Student 16	2.85	3.16	3.00			

Strong Correspondence Between Aggregate Benchmark Rate of Progress and Frequent Monitoring Rate of Progress

- For individual students, the correlation is .54.
- For an aggregate of 10 students, the correlation is .93
- For 30 students, the correlation is .96
- So, a normative context based on aggregate benchmark rates of progress will be directly applicable to frequent monitoring rates of progress, especially when 30 or more students are aggregated.

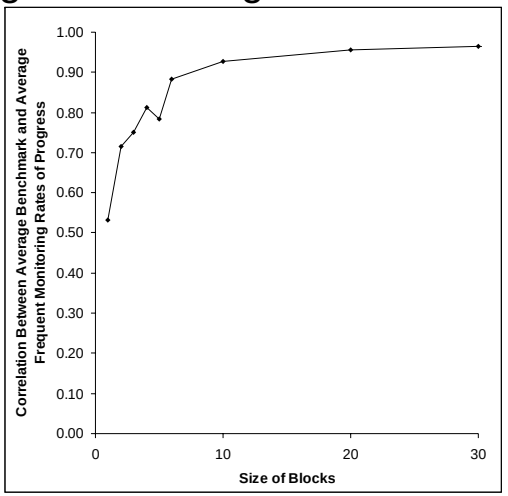


Table 23

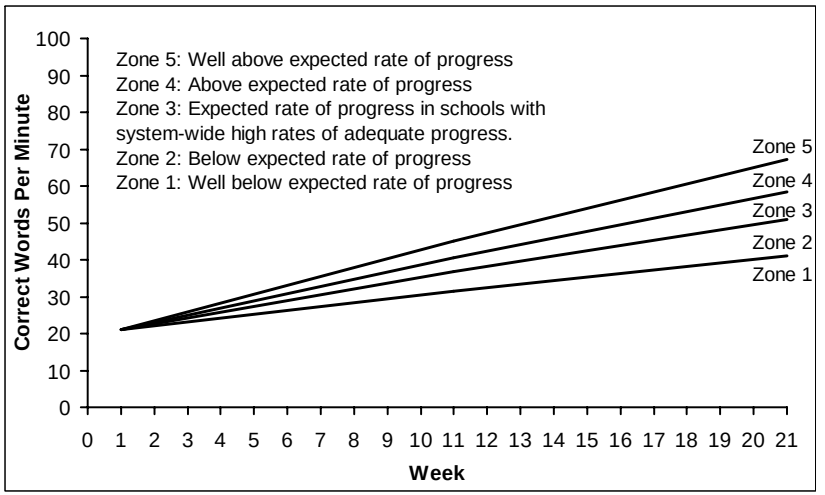
Zones of Growth by Level of Initial ORF Score in Beginning of Year Second Grade to Middle of the Year for Schools with 40 or More Students with High Rates of Adequate Progress for All Three Tiers (Conditional Probability of Intensive Reaching Strategic or Benchmark >= 23 and Conditional Probability of Strategic Reaching Benchmark >= 54 and Conditional Probability of Barely Benchmark Staying at Benchmark >= 95)

		BOY - MOY growth percentile			
BOY ORF	*n*	20th percentile	40th percentile	60th percentile	80th percentile
Intensive					
0 to 5	934	0.11	0.33	0.56	0.98
6 to 15	3145	0.40	0.70	1.05	1.53
16 to 25	6270	0.95	1.43	1.78	2.20
Strategic					
26 to 34	7862	1.30	1.73	2.06	2.43
35 to 43	7415	1.50	1.83	2.11	2.50

Using Zones of Growth Normative Context to Graph Individual Zones

- 1. Start with the student's BOY level of skills. If multiple assessments are given to verify need for support, the median score would generally be a good estimate of initial skills.
- 2. Identify the band of initial performance in the Zones of Growth Norms table.
- 3. Count out 10 weeks from the initial assessment.
- 4. Multiply the growth rates by 10 (move the decimal 1 place to the right) and add to initial skill level.
- 5. Plot the points and use a ruler to draw lines dividing the zones of growth.

Zones of Progress for a student with 21 words correct at Beginning of Year Second Grade



Conclusions and Discussion

- This work furthers our understanding of the reliability of slope based on frequent progress monitoring as an estimate of the rate of progress and a normative context for those rates of progress.
- Many questions remain to be answered.
  - Which normative context is the most defensible and interpretable?
  - Are there other ways of estimating rates of progress that should be considered?
    - Bayesian estimates?
    - Multi-level estimates?
    - Focused point estimates (i.e., get a very precise estimate of student skills at week 10, for example).

What about meaningful rates of progress? What about generally effective instruction?

Evidentiary Considerations for the Educational Decisions Required for Response to Intervention Models

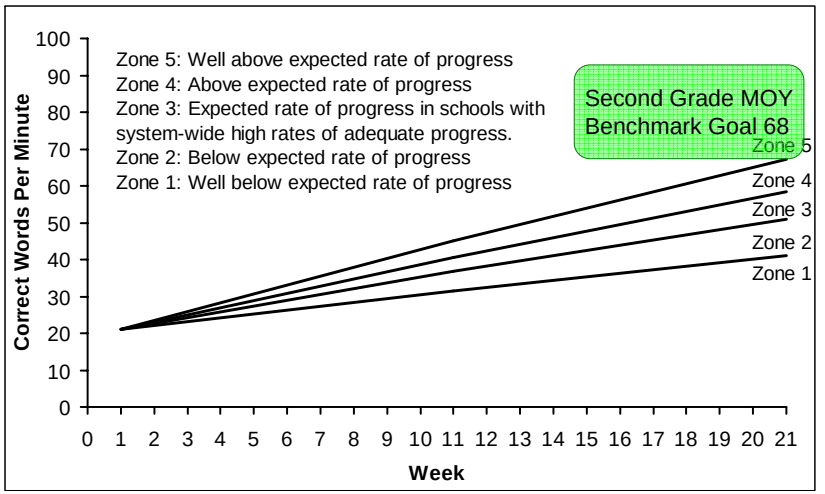
	Evidentiary Consideration		
		Normative	
Educational Decision	Reliable	Context	Meaningful
1. Is the student making adequate year-to-year progress?	X	X	X
2. Is the student receiving generally effective instruction?	?	?	?
3. Is the student making adequate week-to-week progress?	+/-	+/-	+/-

Note. X = generally strong and persuasive evidence. ? = level of evidence is unestablished. +/- = emerging evidence base.

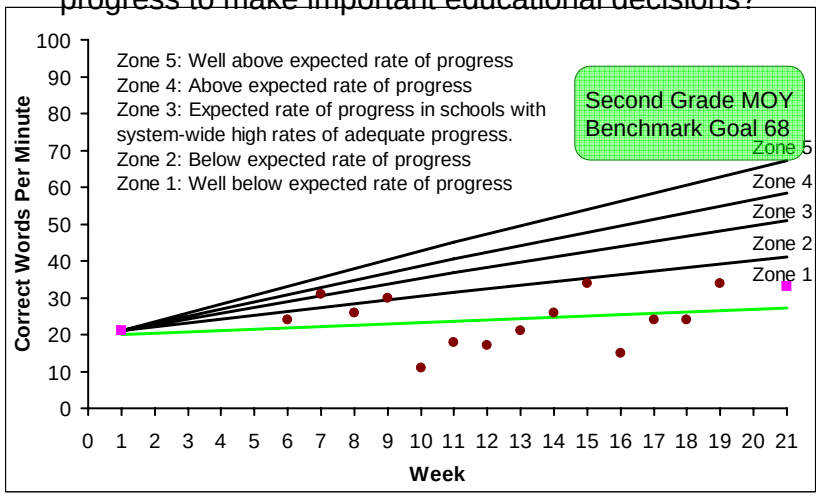
Shouldn't we just wait until the evidence base is complete before using RTI for eligibility decisions?

- Of course, the alternative is to keep using an ability-achievement discrepancy to identify learning disability – there is substantial research on the approach.
  - No evidence that an ability-achievement discrepancy is educationally meaningful.
  - Evidence that an ability-achievement discrepancy does not correspond well to the decisions educators make in practice.
- Or we could rely on individual judgment: "I know them when I see them".
- Or we could suspend eligibility decisions until the scientific basis is completely established.

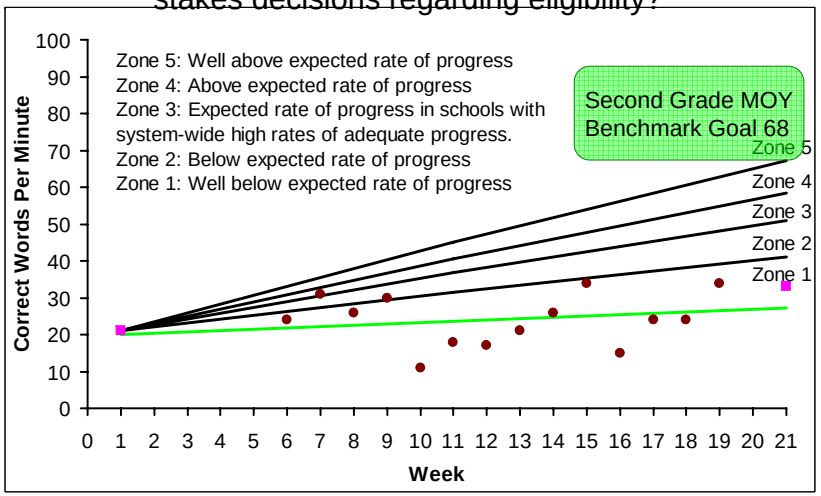
Are zones of progress meaningful? What rate of progress is required to have an important impact on SAT-10 or the state assessment?



We can have more confidence in the Zone lines (r = .96+) than in the individual rate of progress (r = .75).
Do we have enough confidence in the individual rate of progress to make important educational decisions?



How many weeks are needed to have adequate confidence to make low stakes decisions to maximize learning?
How many weeks are needed to make defensible high stakes decisions regarding eligibility?



Thoughts for the Future

- Basing important individual educational decisions like eligibility on ordinary least square slope estimates seems fraught with challenges.
  - Even under pretty good conditions, 14 frequent monitoring + 2 benchmark, Intensive and strategic students, reasonably well-behave progress monitoring data, the reliability of the slope estimate (.75) does not inspire confidence. Of course, there currently are no generally accepted standards for the reliability of slope. But still….
  - The differences between zones of growth over one semester are not large. Above the 40th percentile is not very far from below the 20th percentile.
- Perhaps a different approach to evaluating progress should be considered.

Table 14

School Based Percentile Ranks for the Beginning of the Year to the Middle of the Year of 2nd Grade and Schools with More Than 40 Students

	Initial Status - BOY			Adequate Progress					Outcome - MOY		
School Percentile	Percent Intensive	Percent Strategic	Percent Benchmark	Intensive Adequate Progress	Intensive Exceptional Progress	Strategic Adequate Progress	Barely Benchmark Adequate Progress	Percent Adequate Progress	Percent Deficient	Percent Emerging	Percent Established
1	0	7	19	0	0	0	50	28	1	2	24
5	3	12	28	0	0	11	67	39	4	5	36
10	6	15	34	0	0	17	75	45	7	7	42
15	7	17	38	0	0	21	80	49	9	8	46
20	9	18	41	0	0	25	83	53	10	9	50
25	10	20	43	5	0	29	85	56	12	10	53
30	12	21	46	7	0	31	88	59	13	11	56
35	13	22	48	8	0	33	89	62	15	12	59
40	15	23	51	10	0	37	90	64	16	13	61
45	16	24	53	13	0	40	92	66	18	13	63
50	18	25	55	14	0	43	93	68	20	14	65
55	19	26	57	17	0	45	94	70	21	15	68
60	21	27	60	18	0	48	95	72	23	16	70
65	23	28	62	20	0	50	97	74	25	16	72
70	25	29	64	23	3	54	100	76	27	17	74
75	27	30	67	25	5	57	100	78	30	18	76
80	30	32	69	29	7	60	100	80	32	19	78
85	33	33	73	33	10	64	100	83	36	20	81
90	38	35	76	40	13	69	100	86	40	22	84
95	44	38	82	50	20	77	100	90	47	24	88
99	57	44	89	73	38	90	100	94	60	29	94

Note. Based on 6958 schools with 78176 students with beginning of year second grade ORF scores and middle of year second grade ORF scores.

Table 23

Zones of Growth by Level of Initial ORF Score in Beginning of Year Second Grade to Middle of the Year for Schools with 40 or More Students with High Rates of Adequate Progress for All Three Tiers (Conditional Probability of Intensive Reaching Strategic or Benchmark >= 23 and Conditional Probability of Strategic Reaching Benchmark >= 54 and Conditional Probability of Barely Benchmark Staying at Benchmark >= 95)

		BOY - MOY growth percentile			
BOY ORF	*n*	20th percentile	40th percentile	60th percentile	80th percentile
Intensive					
0 to 5	934	0.11	0.33	0.56	0.98
6 to 15	3145	0.40	0.70	1.05	1.53
16 to 25	6270	0.95	1.43	1.78	2.20
Strategic					
26 to 34	7862	1.30	1.73	2.06	2.43
35 to 43	7415	1.50	1.83	2.11	2.50
Benchmark					
44 to 53	7578	1.48	1.80	2.11	2.53
54 to 63	7263	1.35	1.73	2.08	2.53

Note. Based on 63055 students in 783 schools with high rates of adequate progress for intensive, strategic, and barely benchmark students.