

Creating a Metric for Measuring Early Student Literacy Development: A Rasch Analysis of DIBELS Assessment Data

Gina Biancarosa
Anthony S. Bryk



David Kerbow
Stuart Luppescu



National Reading Conference, December 1, 2007

Current study

- ❑ Assessed all students (K-3) attending the 18 participating public schools across 8 states in the Eastern U.S.
- ❑ Tested in fall and spring for 4 years to measure change over time in students' literacy learning using:
 - Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in fall and spring, grades K-2 and fall 3
 - Terra Nova in spring, grades 1-3

Low Income	40.1%
Race/Ethnicity	African-American 15.5%
	Latino 6.9%
	Asian 4.3%
	Native American 0.5%
	White 72.8%
Limited English Proficiency	4.0%

DIBELS and Terra Nova testing

		K-f	K-s	1-f	1-s	2-f	2-s	3-f	3-s
DIBELS									
	Initial sound fluency	X	X						
	Letter name fluency	X	X	X					
	Phonemic segmentation fluency	X	X	X	X				
	Nonsense word fluency		X	X	X				
	Oral reading fluency			X	X	X	X	X	
Terra Nova					X		X		X

Challenges with the assessments

- Terra Nova inappropriate for K
- DIBELS
 - comprehension subtest has poor reliability
 - different subtests used at different times
 - no continuous scale linking subtests or forms
- No previous linking studies between the two

Scaling early literacy measures

- ❑ Raw scores are **not** measures
- ❑ The meaning of scoring one point more or less depends on the difficulty of the item
- ❑ Measures use a truly *interval* scale: Equal differences on scale imply equal differences on the trait measured at any level
 - True of distance or temperature measures,
 - Untested for DIBELS early literacy raw scores
- ❑ Problematic for research (Paris, 2005)
 - Floor and ceiling effects
 - Skewed distributions
 - Violations of underlying assumptions of parametric statistics: normality, homoscedasticity



Objectives

- To equate the subtests and forms of the DIBELS and create a continuous literacy ability metric for use in further analyses
- Ultimately to equate Terra Nova with DIBELS (stay tuned)

DIBELS: raw scores mean different things

DIBELS subtests all measure fluency in component skills of literacy, but do so in different ways

		K-f	K-s	1-f	1-s	2-f	2-s	3-f	3-s
DIBELS									
	Initial sound fluency	ISF: Unlimited time, but time in seconds weights score							
	Letter name fluency	LNF: Timed, 60 seconds							
	Phonemic segmentation fluency	PSF: Timed, 60 seconds Partial credit given							
	Nonsense word fluency	NWF: Timed, 60 seconds Partial credit given							
	Oral reading fluency	ORF: Timed, 60 seconds				X	X	X	

DIBELS risk benchmark results

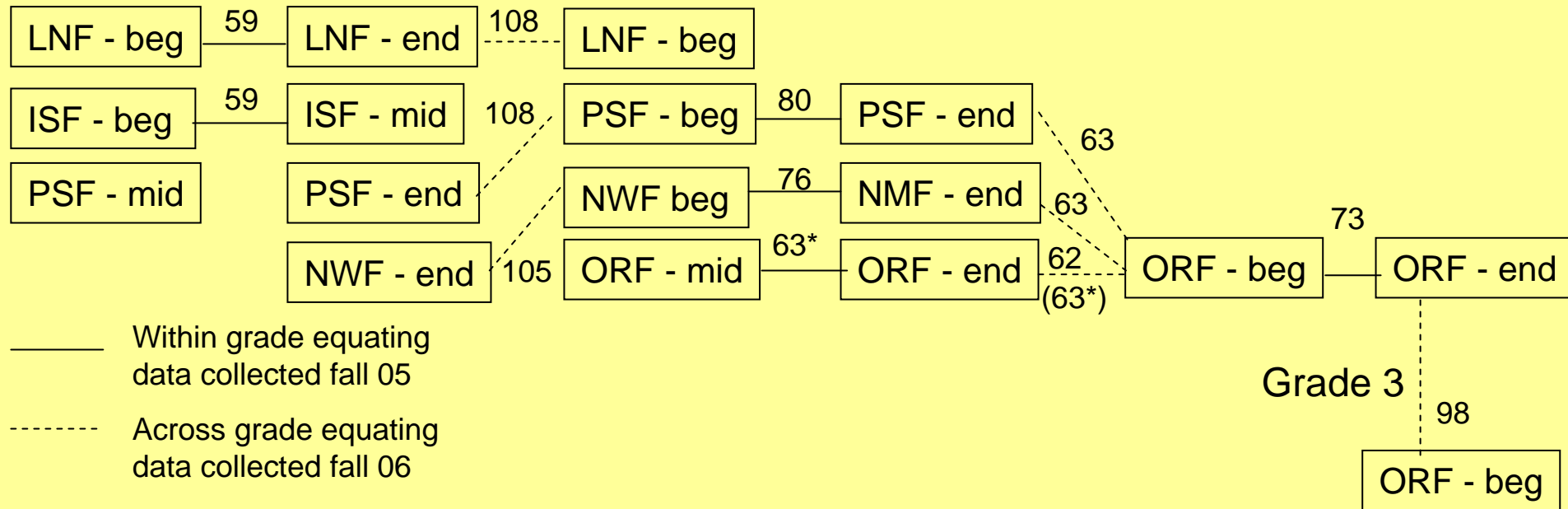
			Year of Study		
			Year 1		Year 2
			Fall	Spring	Fall
Grade	K	Intensive	10%	—	
		Strategic	35%	—	
		Benchmark	55%	—	
	1	Intensive	13%	—	
		Strategic	27%	—	
		Benchmark	60%	—	
	2	Intensive	13%	—	
		Strategic	24%	—	
		Benchmark	62%	—	

Equating design

Kindergarten

Grade 1

Grade 2



Equating procedure

1. Calibrated all items in each vertical sample
2. Calibrated equating items that 10% or more of the sample got correct
3. Calculated equating constants
 1. Calculated difference between forms for equating sample
 2. Calculated difference between forms (using common subtests only) for vertical sample
 3. Calculated difference between these differences to yield the final equating constant
4. Added constants cumulatively (within K, between K-1, within 1, between 1-2, etc.)

Equating results

Testing form/period	Equating constant (difference in difficulty from prior form)	Total difference in difficulty from K fall
K fall	0	0
K spring	+0.58	0.58
1 fall	+2.73	3.31
1 spring	+0.27	3.58
2 fall	+4.82	8.40
2 spring	-1.26	7.14
3 fall	+2.83	9.97

Precision-weighted means from the DIBELS scaling

		Year of Study					
		Year 1		Year 2		Year 3	
		Fall	Spring	Fall	Spring	Fall	Spring
Grade	K	-1.78	-0.08	-2.24	0.39	-2.04	0.15
	1	0.01	0.71	0.20	1.05	0.21	1.53
	2	1.42	2.66	0.68	2.54	1.35	3.10
	3	2.96		2.81		2.98	

Precision-weighted means from the DIBELS scaling

		Year of Study					
		Year 1		Year 2		Year 3	
		Fall	Spring	Fall	Spring	Fall	Spring
Grade	K	-1.78	-0.08	-2.24	0.39	-2.04	0.15
	1	0.01	0.71	0.20	1.05	0.21	1.53

- ▣ *Depending on form difficulty*, Kindergarteners and 1st graders with **-2.00 logits** likely can
 - Answer correctly 11-15 of the 16 Initial Sound Fluency items
 - Name 20-30 letters in a minute
- ▣ But are unlikely able to
 - Segment words phonemically
 - Sound out nonsense words
 - Read > a few first-grade level words per minute

Precision-weighted means from the DIBELS scaling

		Year of Study					
		Year 1		Year 2		Year 3	
		Fall	Spring	Fall	Spring	Fall	Spring
Grade	K	-1.78	-0.08	-2.24	0.39	-2.04	0.15
	1	0.01	0.71	0.20	1.05	0.21	1.53

- *Depending on form difficulty, Kindergarteners and 1st graders with **0.00 logits** likely can*
 - Answer the Initial Sound Fluency items correctly
 - Name 41-48 letters per minute
 - Phonemically segment at least partially 10-15 Kindergarten level words
 - Sound out some nonsense words
 - Read about 13-18 first-grade level words per minute

Precision-weighted means from the DIBELS scaling

		Year of Study					
		Year 1		Year 2		Year 3	
		Fall	Spring	Fall	Spring	Fall	Spring
Grade	K	-1.78	-0.08	-2.24	0.39	-2.04	0.15
	1	0.01	0.71	0.20	1.05	0.21	1.53
	2	1.42	2.66	0.68	2.54	1.35	3.10
	3	2.96		2.81		2.98	

□ Depending on form difficulty, 1st-3rd graders with a **+3.00 logits** likely have mastered

- Initial sounds
- Letter naming
- Phonemic segmentation

□ And can read

- About 29 nonsense words per minute
- About 99 1st-grade wpm
- About 94 2nd-grade wpm
- About 86 3rd-grade wpm

Implications

- DIBELS can be used to track longitudinal change in children's literacy development
 - BUT cut-scores give a stilted picture of this development
 - WHEREAS scaling DIBELS
 - accounts for form and grade-level changes in difficulty and
 - allows for a better estimation of ability

Thank you!