



USAID
FROM THE AMERICAN PEOPLE



WORLD BANK GROUP

Beyond Access: Effective Reading for All and the Early Grade Reading Assessment

Amber Gove, PhD
RTI International

World Bank Institute

April 30, 2009

Contact: agove@rti.org

www.eddataglobal.org

About the presentation

- This presentation was prepared for a World Bank Institute seminar in Washington, DC, on April 30, 2009.
- Much of the work described here is being carried out under the U.S. Agency for International Development's (USAID's) EdData II project, led by RTI International. Early Grade Reading Assessment (EGRA) activities take place under EdData II Task Order Number 3, EHC-E-00-03-00004-00.
- The World Bank also contributes funding for EGRA-related efforts in various countries.

Outline

1. Rationale
2. Development
3. Pilot testing
4. Results and implications

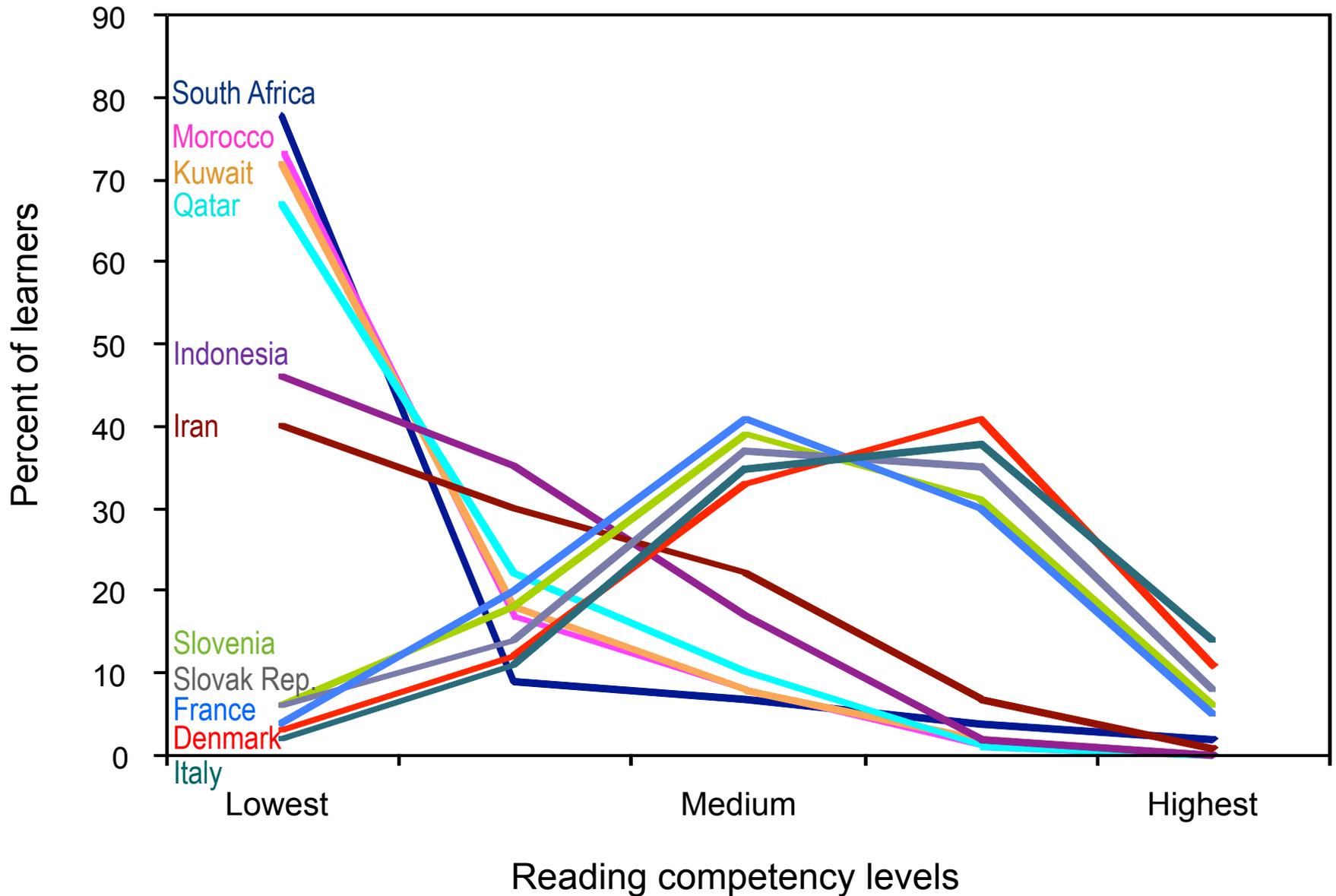
Outline

1. Rationale
2. Development
3. Pilot testing
4. Results and implications

Quality issues

- What are the big international goals?
- How do low-income countries compare to high-income countries?
 - LI to HI ratio
 - Gross primary enrollment: 95%
 - Net primary enrollment: 80%
 - Gender parity net enrollment rate (NER): 94%
 - Completion: 58%
 - Learning achievement: Median LI = 3rd percentile of HI or lower

PIRLS 2006 Results



Why focus on early grade reading?

Early Grade

+

Reading =

- intervene early,
- intervene on reading,
- have some way to assess orally

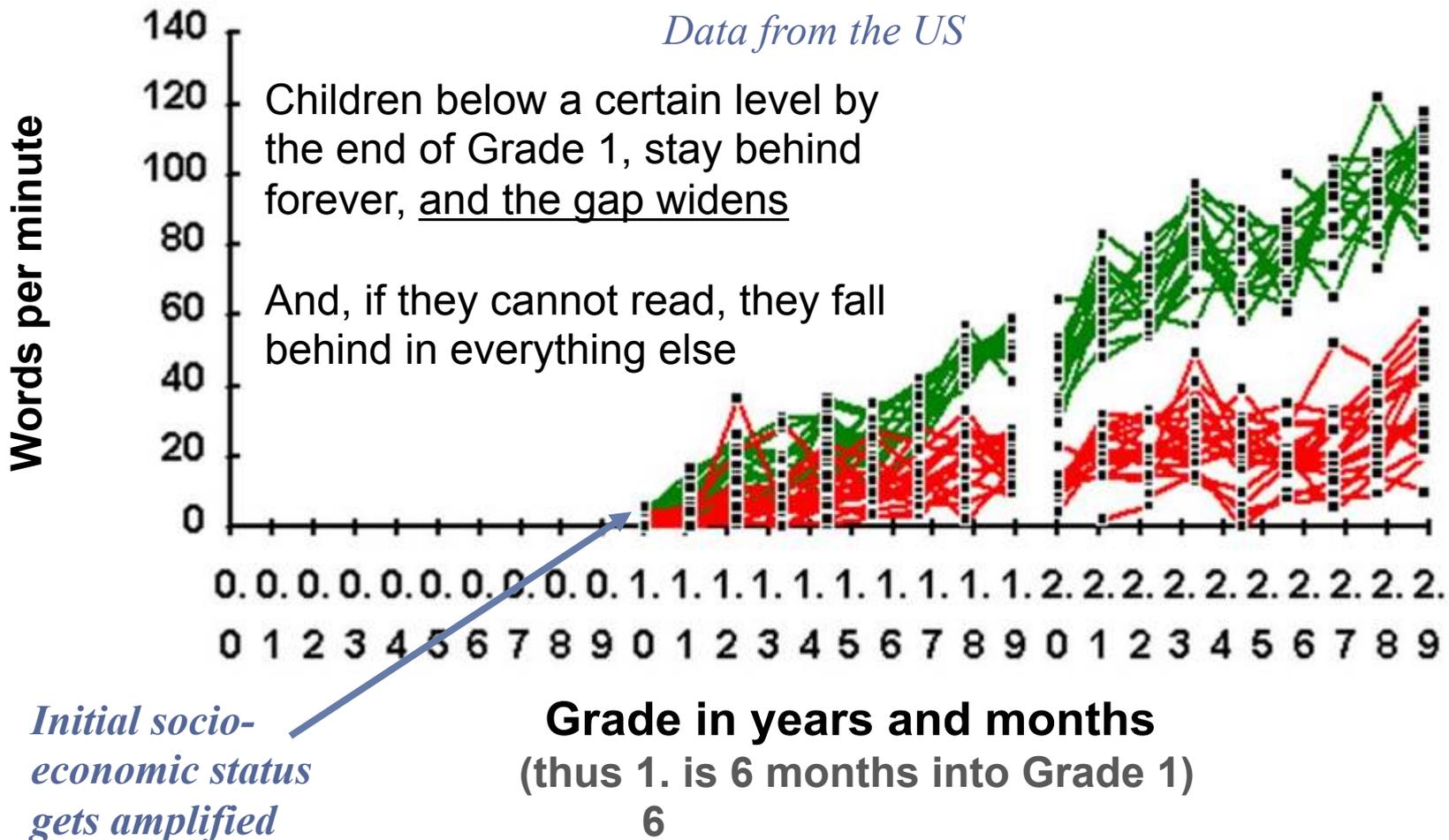
Let's see if we can motivate those conclusions

Why early?

“For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath.”

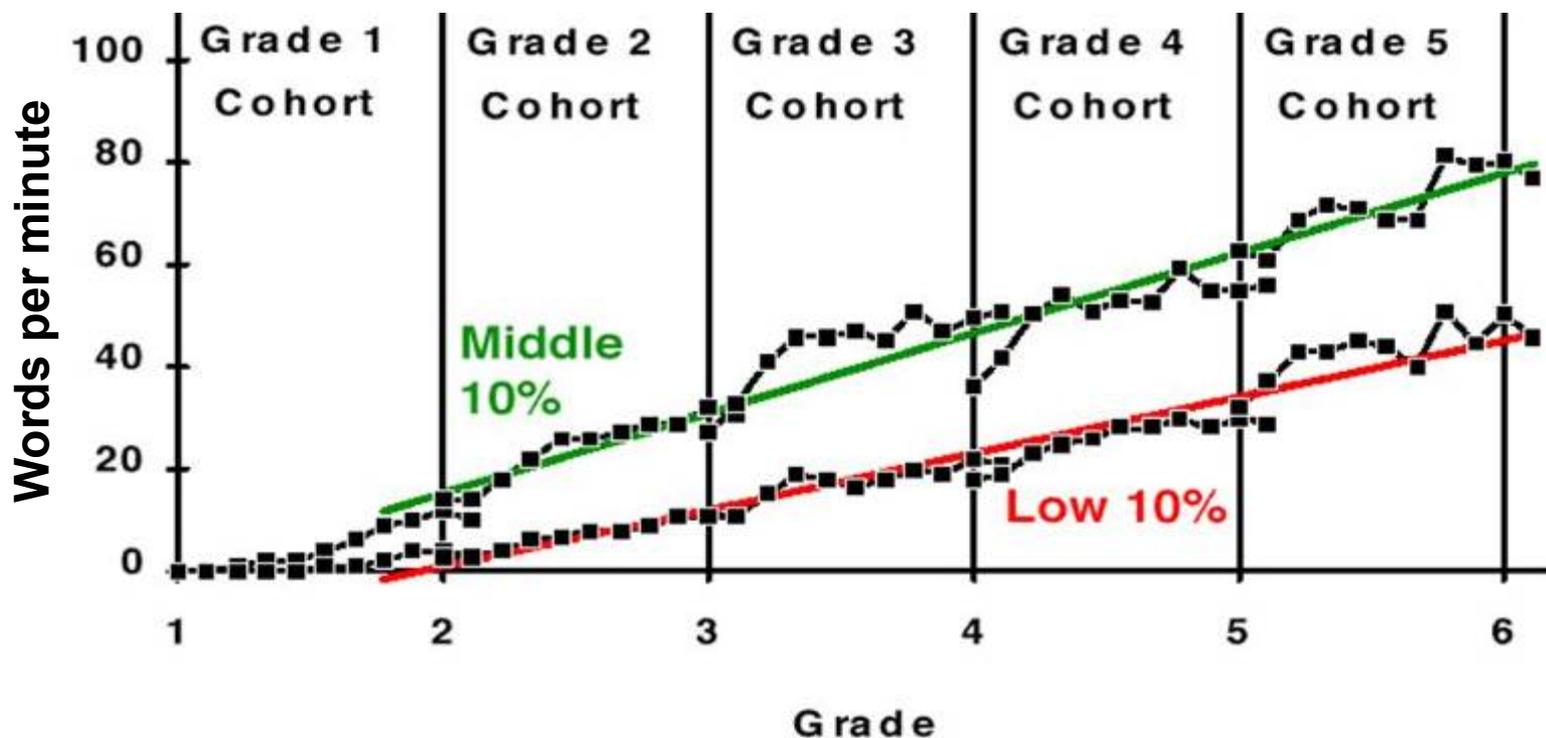
Matthew Effect?

Why early? Matthew Effect in reading



Why early?

Reading Trajectories of Low and Middle Readers



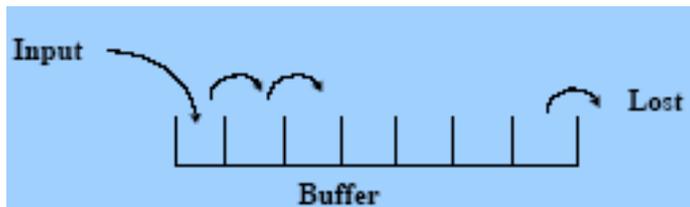
Why reading?

- No, it is not “the only thing that matters”
- But it is a good one to start with
 - It is a (*the?*) foundational skill -
Hard to imagine anything else going well if children can't read well and soon
 - It can be used as a marker -
Hard to imagine a good school that can't teach children to read; if children are not reading, the school (district, country) needs serious help

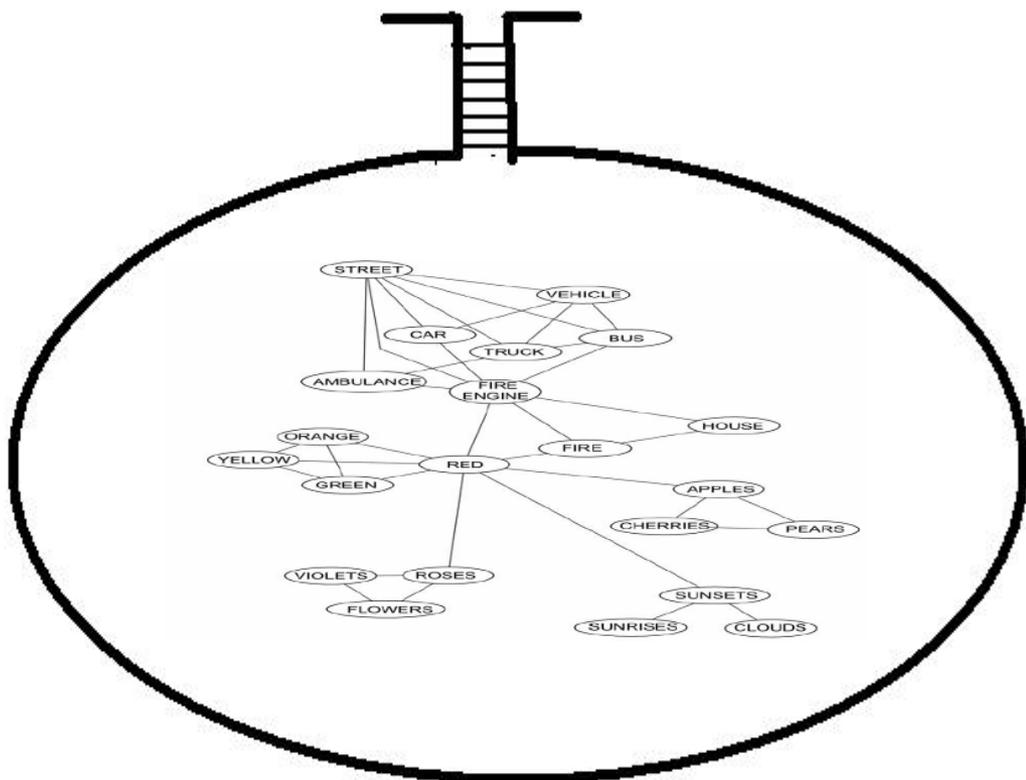
Why oral reading?

- Oral reading seems to be good predictor (see literature)
- Students frequently bottom out (floor-effect problems) on paper-and-pencil tests
- Elements of oral reading are in accord with curricular frameworks but frequently there are no specific (teacher-level) guidelines on how to assess

Why timed oral reading?



- From brain research we know short-term memory is crucial for reading comprehension
- Short-term memory can hold about 7 items for 12 seconds
- Fluency and accuracy are related to comprehension



Oral reading predictive power

Examples:

- Wilson (2005): 0.74 correlation between oral reading measures and broader cognitive achievement in Arizona.
- Fuchs, Fuchs, Hosp, & Jenkins (2001): survey and explain the rather high correlations between oral reading fluency (ORF) and a large variety of other tests.
- Vander Meer, Lentz, & Sorrels (2005): 96% of children judged to be at risk using oral reading turned out to be “nonproficient” in the Ohio’s more comprehensive test, while of those classified as “low risk” using oral reading fluency, 72% were classified as proficient using a more comprehensive test.
- Shaw & Shaw (2002): found similar results for the relationship between simple measures of oral fluency and deeper state-wide measures of reading in Colorado.
- Fuchs, Fuchs, & Maxwell (1988): correlation of 0.91 between oral reading fluency and other comprehensive tests.

Oral reading predictive power (*cont'd*)

- Juel (1988): “The probability of remaining a poor reader at the end of fourth grade, given a child was a poor reader at the end of first grade, was .88 the probability of remaining an average reader in fourth grade, given an average reading ability in first grade, was .87.”
- Schilling, Carlisle, Scott, & Zeng (2007): 80% of at-risk with ORF turned out to be in bottom quartile with Michigan’s own reading test at end of grade.
- Wood (2006): found significant correlations between oral reading fluency and later curriculum-based tests, and found that oral reading fluency adds explanatory value even when other factors are considered.
- Some of these recommend adding comprehension and vocabulary (EGRA does comprehension, not vocabulary), but ORF by itself does a pretty good job.

Outline

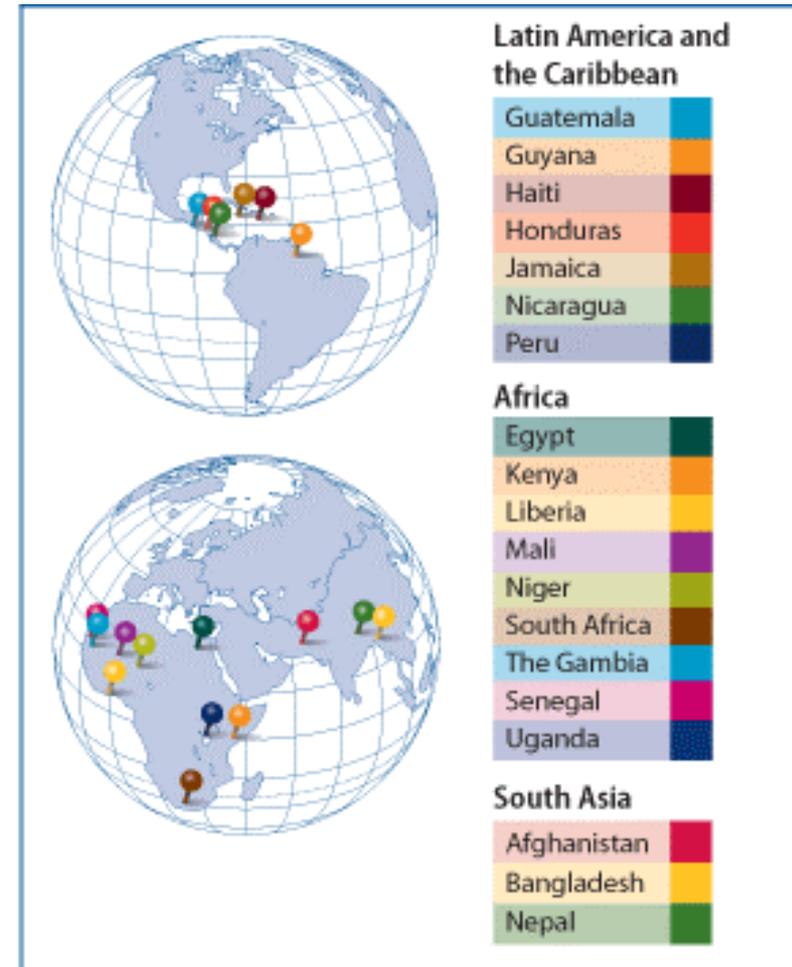
1. Rationale
2. Development
3. Pilot testing
4. Results and implications

Development thus far

- “Organic” process: meets “market test” at each step
- First: informal, small samples, see if it was useful at generating awareness, very little funding
- Attention attracted, community of practice involved
- Some funding to try it a bit more formally
 - USAID funding: validate efforts thus far with expert opinion, try some more applications
 - Call high-level experts meeting, experts validate, suggest increased formality, seriousness of trials
 - World Bank adds some funding, try it in two more international languages, local languages

Development thus far: Gaining momentum

- World Bank financed in 7 countries; USAID in 7 more and growing
- March 2008 workshop: 200 participants from 40 countries
- Colleagues from AED, AIR, Save, IRC, BRAC, Plan, Pratham (and others?) experimenting with EGRA
- 2009 Hewlett Foundation support for work in 9 languages in four countries



Instrument Component	Early Reading Skill	Skill demonstrated by students' ability to:
1. Engagement and Relationship to Print	Orientation to print	Indicate where to begin reading (uppermost left corner)
		Indicate direction of reading within a line (left to right)
		Indicate direction of reading within a page (top to bottom)
2. Letter Naming	Letter recognition	Provide the name of upper- and lower-case letters distributed in random order
3. Phoneme Segmentation	Phonemic awareness	Segment words into 2 to 5 phonemes, counting of phonemes within words
4. Familiar Word Reading	Word reading	Read simple and common one and two syllable words
5. Nonsense Word Decoding	Alphabetic principle	Make grapheme-phoneme correspondences (GPCs) through the reading of simple nonsense words
6. Paragraph Reading and Comprehension Questions	Oral reading fluency	Read a text with little effort and at a sufficient rate
	Reading comprehension	Respond correctly to different type of questions (literal with options, literal and inferential) about they text they have read
7. Listening Comprehension	Listening comprehension	Respond correctly to different type of questions (literal with options, literal and inferential) about they text the enumerator reads to them
8. Dictation	Alphabetic principle	Write, spell and use grammar properly through a dictation exercise

Other possible components for assessment (under exploration)

- Sounds (e.g., ck, oo, nasal vowels—tried in Arabic)
- Vocabulary (especially for children not learning in mother tongue—tried in French/Creole)
- Nonsense word spelling (may work when we have floor-effect problems with traditional phonological awareness (PA) tasks—tried in French/Creole)
- Other texts for assessing comprehension? (in addition to narrative)

Other characteristics

- Many segments timed to 1 minute
 - Meant to measure fluency (critical skill), also more humane, and also more efficient – whole test can be done in less than 15 minutes
- Can be done on sample basis; easy to score
- Can provide essentially instant results for a village
- Applied by assessor (or teacher), one-on-one, not pencil-and-paper in whole class
- Requires training of enumerators (including underlying principles of languages being tested)

Outline

1. Rationale
2. Development
3. Pilot testing
4. Results and implications

Language pilot results



Language(s)	Grades	Schools	Students
English	1, 2, 3	40	1200
French	1, 2, 3	36	501
Wolof	1, 3	36	186
Spanish	1, 2, 3	41	1924
Miskitu	1, 2, 3	6	282
Arabic	1, 2, 3	3	95

Summary of EGRA results: English



- At standard of 45 correct words per minute (cwpm) for paragraph reading, percent of tested students who “can read”:
 - G1: 1.4% (Comp. 96%)
 - G2: 2.6% (Comp. 96%)
 - G3: 6.1% (Comp. 92%)
- Characteristics of “Readers” (N=40) (NR=Nonreaders)
 - 90% have reading books at home (NR: 24%)
 - 57% say mother is literate in English (NR: 22%)
 - 88% attended preschool (NR: 63%)

Summary of EGRA results: French and Wolof

- At standard of 45 cwpm for paragraph reading, percent of tested students who “can read”:
 - G2 French: 11% (Comp. 62%)
 - G3 French: 48% (Comp. 66%)
 - G3 Wolof: 28% (Comp. 77%)
- Characteristics of French “readers” (NR=Nonreaders)
 - 64% have reading books at home (NR: 47%)
 - 85% attended preschool (NR: 77%)
- Characteristics of Wolof “readers” (NR=Nonreaders)
 - 45% have reading books at home (NR: 44%)
- 25 ■ 76% attended preschool (NR: 72%)



Summary of EGRA results: Spanish

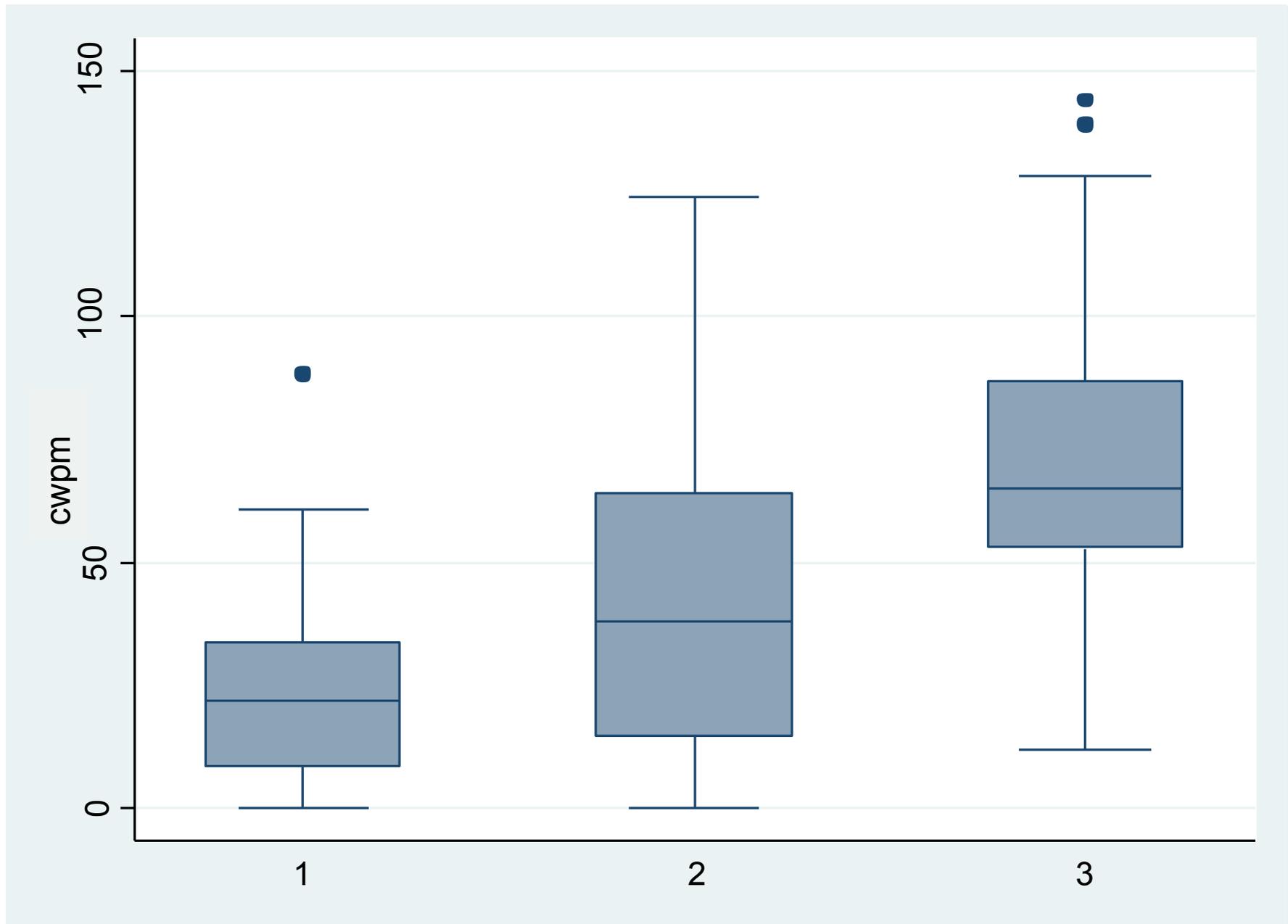


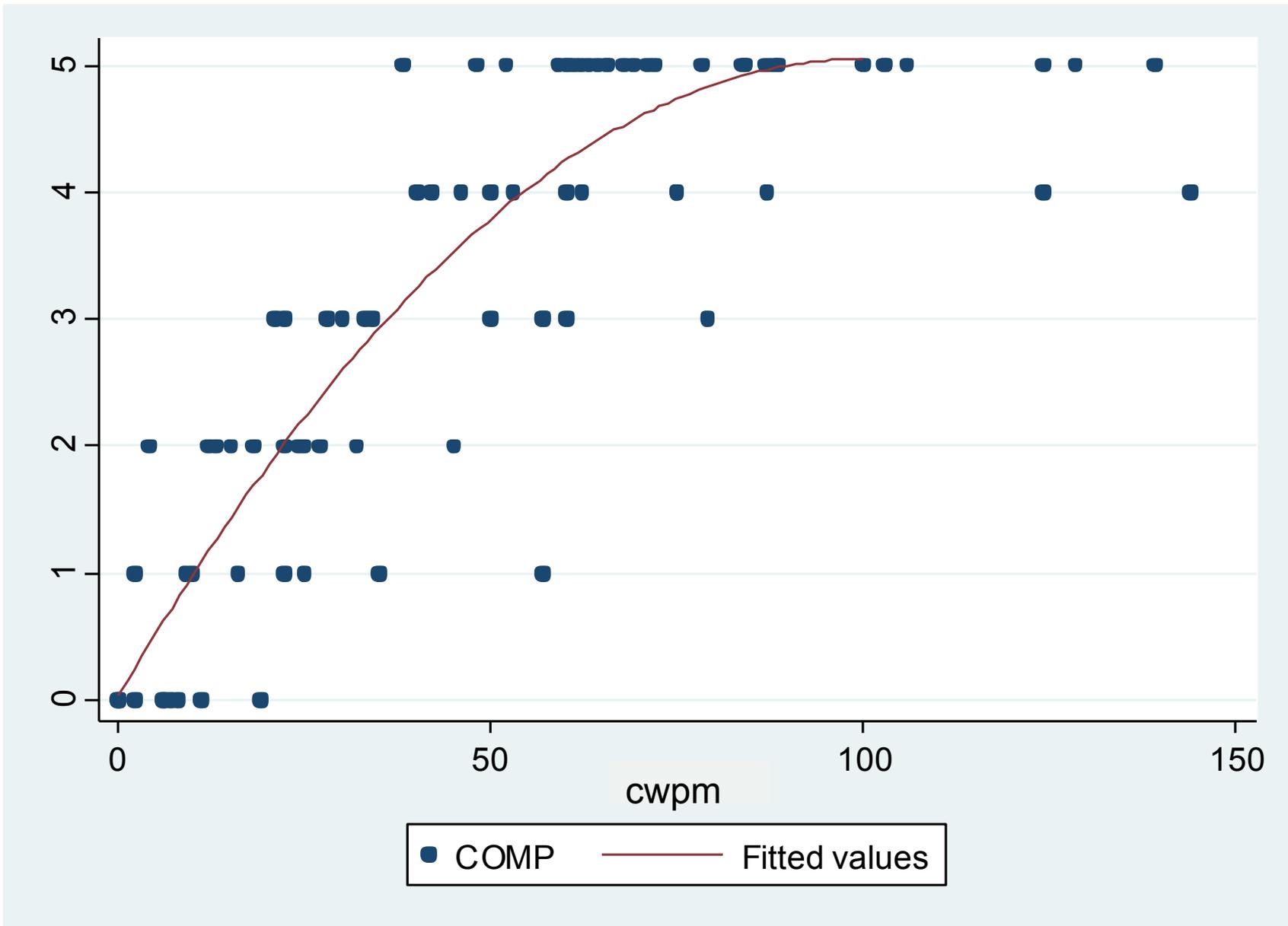
- At standard of 60 cwpm for paragraph reading, percent of tested students who “can read” in Spanish:
 - G1: 17%, Comp, 87%
 - G2: 60%, Comp, 87%
 - G3: 85%, Comp, 87%

Summary of EGRA results: Arabic

- At standard of 45 cwpm for paragraph reading, percent of tested students who “can read” in Arabic:
 - G1 : 12%, Comp, 80%
 - G2 : 41%, Comp, 90%
 - G3 : 80%, Comp, 90%







Floor-effect issues

- Large number of zero scores (higher in those without mother-tongue instruction and/or access to print)

- French: 23%

- Wolof: 36%

- English: 65%

- Spanish: 9%

- Miskitu: 37%

- Arabic: 3%

Outline

1. Rationale
2. Development
3. Pilot testing
4. Results and implications

Results: Reliability and validity

- Reliability—Accuracy and internal consistency
 - Cronbach's alpha: Fr=0.90, Eng=0.87, Sp=0.84, Arabic=0.94
- Validity—Ability of a measure to report outcomes we know to be true
 - Face: Policy makers, parents view EGRA as related to their “sense” of reading
 - Concurrent/External: Requires external measure
 - Predictive: Requires longitudinal data, multiple measures
- Item response theory (IRT) analysis of letters, words, etc., as items

Results: Refining the tool

- Large range on results (e.g. min/max 2 to 120) (good lessons for discussing implications of large standard deviations [SDs])
- Good reliability (alpha= ranging from 0.84 to 0.94)
- Tool discriminates well between performers; showing good grade progression results
- Some (expected) need for refinements; experimentation with new subtests
- Harder than it looks (e.g., lack of word frequency lists in Wolof, Bangla, Miskitu, etc.)
- Arabic pretest demonstrates that principles of alphabetic languages apply even to non-Latin scripts

Possible uses

- Policy awareness and motivation
 - Macro
 - Community-based
- Impact tracking and evaluation
 - Project monitoring
 - Project impact and evaluation
 - System monitoring over time
- Teacher-based assessment
 - (Could link to community-based awareness, accountability?)
 - Motivating and driving instructional practice
 - Already being used that way in various countries

Uses to date

- Revamped approach to teacher professional development and focus on the early grades
- Generated national debate and presidential pledge to have all children reading by the end of grade 2
- Renewed focus on providing quality teacher professional development and instructional materials in local languages
- Spurred development of reading intervention programs and continuous monitoring by teachers

References

- Fuchs, L., Fuchs, D., Hosp, M. K., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*(3), 239-256.
- Fuchs, L.S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education, 9*, 20-28.
- Good, R.H., III, Simmons, D.C., & Smith, S.B. (1998). Effective academic interventions in the United States: Evaluating and enhancing the acquisition of early reading skills. *School of Psychology Review, 27*(1), 45-56.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grade. *Journal of Educational Psychology, 80*(4), 437–447.
- Schilling, S. G., Carlisle, J. F., Scott, S. E., & Zeng, J. (2007). Are fluency measures accurate predictors of reading achievement? *Elementary School Journal, 107*(5), 429-448.

References (cont' d)

- Shaw, R. & Shaw, D. (2002). *DIBELS oral reading fluency-based indicators of third grade reading skills for Colorado State Assessment Program (CSAP)*. Technical report. Eugene, OR: University of Oregon.
- Vander Meer, C. D., Lentz, F. E., & Stollar, S. (2005). *The relationship between oral reading fluency and Ohio proficiency testing in reading*. Technical Report. Eugene, OR: University of Oregon.
- Wilson, J. (2005). *The relationship of Dynamic Indicators of Basic Early Literacy Skills [DIBELS] oral reading fluency to performance on Arizona Instrument to Measure Standard: (AIMS)*. Technical Report. Tempe School District No. 3, Tempe, Arizona.
- Wood, D. E. (2006). Modeling the relationship between oral reading fluency and performance on a statewide reading test. *Educational Assessment*, 11 (2), 85–104.