

**Measuring Implementation in Research & Practice:  
*Developing an Implementation Measure  
for Project Upgrade***

***Carolyn Layzer & Jean Layzer***  
**2009 Annual Meeting of the Child Care Policy Research Consortium,  
October 28-30, 2009 – Washington, D.C.**

**From Insight to Impact – Worldwide**



## Work on this project was funded by:

- U.S. Dept. of Health and Human Services, Administration for Children and Families Child Care Bureau — *Child Care Subsidy Evaluation*, contract #233-01-0015
- Development of original BTL fidelity crosswalk funded by Institute of Education Sciences, U.S. Department of Education, through the Grant R305G04145 to the University of Iowa, Abt Associates Inc. subcontractor (CLIMBERs study).

## Why measure implementation in a randomized experimental evaluation?

- Helps evaluator understand findings (especially absence of impacts) by answering two questions:
  - To what extent was the intervention implemented as planned by members of the treatment group? (*fidelity*)
  - To what extent did members of the comparison group engage in the same types of interactions, use the same materials and strategies, as the members of the treatment group? (*implementation or contamination/crossover*)

## Approaches to measuring implementation (or fidelity, if only treatment group is measured)

- The designer of the intervention develops a measure that can be used to monitor implementation and shape ongoing training and mentoring efforts.
- The evaluator works closely with the designer to develop a measure of implementation.

\* “Vertical” vs. “Horizontal” measures of implementation

# Challenges for Measuring Implementation in Project Upgrade

- Project Upgrade was a rigorous experimental test of three language/literacy interventions for four-year-olds from low-income families in child care settings in Miami-Dade County.
- Each developer/designer used a fidelity measure specific to his/her intervention, applied by coaches/mentors at regular intervals.
- We (the evaluator) observed classroom behaviors, interactions and aspects of the environment related to support for children's language and literacy development across all classrooms in the study.
- OMLIT was used to measure proximal outcomes, in terms of teacher behaviors and interactions that we believed to be necessary mediators of the desired child outcomes.

## Study Findings

- Across all classrooms in the study, OMLIT scores for teachers and the environment were low at baseline.
- Most classrooms in the three treatment groups reached high levels of fidelity to their specific intervention (**as rated by coaches/mentors**) by the end of the study
- All three interventions had significant impacts on teacher and environmental support for children's language and literacy development (OMLIT).
- Two of the three had significant impacts on children's language and pre-literacy skills.
- Impacts were strongest for Spanish-speaking teachers and children.

# Strategy: Using OMLIT Variables/Constructs to Measure Implementation

- For each intervention, take the existing implementation measure, group items conceptually (phonemic awareness, text conventions, etc.) if developer has not done so.
- Identify items from OMLIT battery of measures that reflect the same concepts.
- Assess adequacy of resulting measure.
- Apply measure to OMLIT data to derive (new\*) implementation scores for each treatment classroom and control classroom.

# How it Worked for Breakthrough to Literacy (BTL)

BTL Dimension (Reads & Discusses)	Comment	Dimension on Index
a. BOW displayed prominently in the classroom (such as on an easel, easy to see)	Not captured in OMLIT	n/a
b. Graphic organizers displayed and used in instruction 4-5 times per week	Not captured in OMLIT	n/a
c. All/Most BTL reading support materials visible and accessible (Elephant, BOW posters, Alpha frieze, audio CDs, pupil books, book baskets). Visible & accessible: displayed at eye level of children, used by children.	Not captured in OMLIT	n/a
d. BOW is read aloud daily or as often as possible	Should be observed	1
e-1. Teacher meets with small groups to re-read the BOW.	If every day, should be observed	2, 3
e-2. Teacher uses BTL Teacher Guide to provide comprehension and word study activities for small groups, determined by child reports, performance and observation data.	Can only capture gist (likely 3-4 rating on original fidelity scale)	3
f. Teacher emphasizes vocabulary/ oral language 4-5 days during a typical week.	Should be observed	4, 5
g. Teacher emphasizes comprehension strategies/higher-order thinking skills 4-5 days during a typical week.	Should be observed	3

## After identifying variables...

- Either
  - Assign points to each subscale, total, standardize, combine to form composite index
- or
  - Look at what proportion of teachers implemented each dimension or characteristic/component – This will enable you to see which elements teachers were less likely to implement

# How it Worked for Ready, Set, Leap! (RSL)

RSL Dimension (Curriculum)	Comment	Dim. on Index
1. RSL! Materials are set up throughout the classroom	Not measured	n/a
2. The teacher has implemented a system that ensures that children appropriately use the RSL technology components during learning center time and have learned to take care of it (please document teachers system on comment sheet)	Not measured	n/a
3. There is evidence in the classroom of the use of RSL! Materials	Text comments	11
4. The teacher follows the daily schedule	Not measured	n/a
5. Opportunities for extended uses of language, (activities, materials and displays) are coordinated with RSL! goals in each unit and its current theme		12
6. The teacher follows a lesson plan from the “Teacher’s Manual” or the one provided by the project coordinator	Not measured	n/a
7. Teacher encourages the use of the LeapMat to review the letters learned in the current lesson, and follows suggested activities in the “Teacher’s Manual”	Text comments	13
8. Teachers integrates the LeapPad, the correspondent book, and activities (“Teacher’s Manual”) with the current lesson	Text comments	14
9. Teachers encourage the use of the LeapDesk , and models its use to review letters and numbers as suggested in the Manual	Text comments	15
10. Teacher reads the suggested book for the lesson (Read-Aloud Library) , and implements suggested activities in the “Teacher’s Resource Guide”	RAP	16
11. The teacher integrates “ Little Ears” CD or tape to the lesson (note in comments section if music is simply played vs. actively integrated)	Not measured	n/a
12. The teacher uses the big books and Read Aloud Library for full-group and small group activities according to the lesson	RAP	17

# How it Worked for Building Early Language and Literacy (BELL)

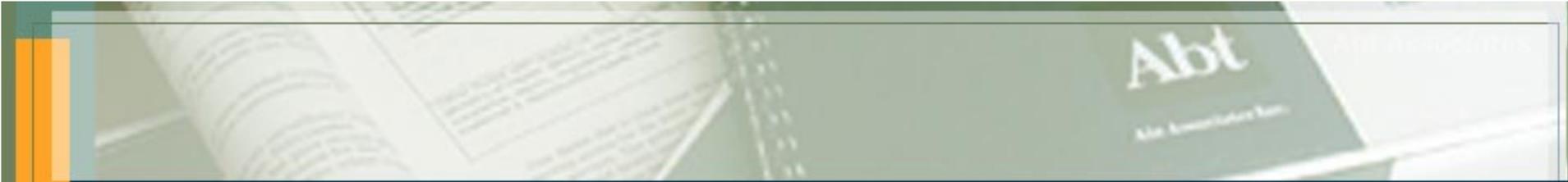
BELL Dimension	Comment
<i>Oral Language Development</i>	
1. Teacher discusses related concepts & vocabulary	Any vocabulary activity—too general
2. Teacher asks questions related to rhyme (meaning)	Comprehension of text—too general
3. Teacher & students recite or sing rhyme	Singing—too general
4. Teacher & students dramatize rhyme	Dramatization related to text
<i>Word Awareness</i>	
5. Students push blocks for each word	Too specific
6. students clap for each word	Too specific
<i>Rhyming Words</i>	
7. use rhyming objects or pictures	Sounds with objects/pictures
8. make a rhyming picture web	Too specific
9. play “thumbs-up/thumbs-down” game	Too specific
<i>Compound Words/Syllables</i>	
10. Use 2-part pictures or puzzles	Too specific
11. feel syllables with hand under chin	Too specific
12. use connecting links or blocks to show awareness of syllables	Too specific
13. use fists to “break it and make it”	Too specific

## Next Steps

- Apply the newly-derived measures for BTL and RSL to:
  - Produce fidelity/implementation scores for classrooms in each of the 2 treatment groups and for the control group classrooms
  - Explore/describe features of each treatment that teachers were more/less likely to implement
- Investigate, with great caution, links between:
  - Implementation levels and characteristics of teachers
  - Implementation levels and outcomes for children.

## Caveats

- Important aspects of implementation, such as intensity or exposure (e.g., number of times you need to use a strategy) or quality of the interaction or behavior (as opposed to whether it was observed) not always made explicit in developer's measure.
- Weight / Relative importance of components not explicit (and generally hard to figure out?)
- Can only measure what designer explicitly stipulates.



**Abt Associates Inc.**