Abstract Title Page

Not included in page count.

Title: *Foundations of Science Literacy: Using instruction-embedded formative assessment to strengthen the relation between gains in teacher pedagogical content knowledge and children's scientific thinking.*

Authors and Affiliations:

Nancy Clark-Chiarelli, Jess Gropen, Janna Fuccillo, Cindy Hoisington
Education Development Center

Abstract Body

Limit 4 pages single-spaced.

Background / Context:

The need for a scientifically-literate American population has gained great prominence on the educational landscape of priorities over the past decade; young children will be increasingly exposed to the STEM fields and encouraged to excel in these areas (The White House, 2009). While experience and research suggest that teachers' science knowledge is predictive of children's science learning (Loucks-Horsley, Love, Stiles, Mundry, & Hewson, 2003), many early childhood teachers are not ready to engage children in rich science experiences that lay the groundwork for later success. Instead, teachers often rely on "spur of the moment" planning or provide activities that are loosely connected to a theme (Bowman, Donovan, & Burns, 2001). Even quality programs tend to emphasize language and social development, at the expense of science learning (Smith & Dickinson, 1994). These challenges are exacerbated by some teachers' apprehension towards teaching science based on uncertainty of both the science content and pedagogy. When science instruction does occur, it usually focuses either on the recall of facts with insufficient attention to student understanding or, alternatively, on activities that are only loosely related to conceptual goals. Neither approach promotes teacher-student interaction that furthers science learning (Duschl, Schweingruber, & Shouse, 2007).

Professional development is key to assuring that early childhood teachers provide children with cognitively-challenging early learning experiences (Bowman, et al., 2001; Dwyer, Chait, & McKee, 2000; Espinosa, 2002; Helburn & Bergmann, 2002). Unfortunately, few models of professional development build preschool teachers' skills and knowledge in an ongoing way and provide access to content knowledge (Barnett, 2003; Whitebook, 2003). Recent research comparing different types of professional development for science teaching suggests that the most effective models are a hybrid, combining professional development with curriculum in ways that lead to intentional and informed use of curricular materials (Penuel & Gallagher, 2008). In addition, effective teacher professional development should not only focus on developing teachers' science content knowledge, but also their *pedagogical* content knowledge related to children's early science development (Ball, 2000; Shulman, 1987).

Purpose / Objective / Research Question / Focus of Study:

We are responding to the critical need for empirical evidence on effective strategies to improve science instruction in preschool. Focusing on the Head Start community, *Foundations of Science Literacy (FSL)* is a credit-bearing professional development course that directly addresses the achievement gap in early science education. The program not only addresses an urgent need, it also integrates the resources, structure, and support that preschool teachers need to improve early science learning and teaching. Based on many years of experience, we have learned that episodic workshops, offered without a sound curriculum or credit, do little to change teachers' classroom practice. In sharp contrast, the great promise of *FSL* is that it includes several features that create a comprehensive approach. One new feature, under current development, is the use of embedded formative assessments to guide children's science learning. In principle, formative assessment is a powerful tool for helping teachers understand children's current knowledge and support their reasoning, reflection, and conceptual development.

In this paper, we draw on two current IES-funded projects in order to address two research questions. First, based on data from an efficacy study nearing completion, we assess whether *FSL without* embedded formative assessments leads to gains in preschool children's understanding of science and scientific thinking as mediated by improvement of preschool teachers' pedagogical content knowledge and quality science instruction within the classroom. Second, drawing on insights from a new development project, we describe how instruction-embedded formative assessment can be used to *strengthen and clarify* the relation between gains in teacher pedagogical content knowledge and gains in children's scientific thinking. In particular, we argue that formative assessment is critical in strengthening this relation, as it provides much-needed structure for preschool teachers to consolidate their pedagogical content knowledge in science, and translate it into effective instruction. In addition, unlike one-shot summative assessment, formative assessment is also critical in clarifying this relation, as it provides a means of tracking and relating the co-occurring changes in teachers' instruction and children's scientific thinking at various points during the implementation of the program.

Setting:
For the efficacy study, *FSL* was implemented in the greater New York City area. Implementation and research for Cohort 1 of the study were conducted during the 2009-2010 school year. (As available, we will also discuss data for Cohort 2, conducted during the 2011-2012 school year.) For the development project, *FSL* is being implemented in Hartford CT.

Population / Participants / Subjects:
Our population includes preschool teachers and a sample of children in their classrooms. The Cohort 1 analytic sample in the efficacy study is comprised of 40 intervention classrooms and 32 control classrooms, with 436 children (270 in intervention classrooms and 186 in control). The teachers and children participating in the development project are drawn from three early childhood groups in Hartford: Community Renewal Teach Head Start, Harford Public Schools, and the Office for Young Children that oversees Hartford's School Readiness Program.

Intervention / Program / Practice:
FSL as assessed in the efficacy study consists of three components: 1) instructional, face-to-face sessions that build teachers' pedagogical content knowledge in physical science; 2) a coaching component that supports teachers as they master content and methods of inquiry in physical science; and 3) a unit of the Young Scientist Series, a unique preschool science curriculum for 4-year-olds in widespread use and with recognition from national science education organizations. In the development project, *FSL* is being expanded to include additional science domains (life and earth sciences; engineering) as well as a toolkit of formative assessments designed to help teachers assess children's scientific understanding and thinking and plan effective instruction.

Research Design (Efficacy Study):
We used a randomized controlled trial (RCT) design with a total sample of 78 preschool classrooms, with 40 classrooms assigned to the intervention group and 32 classrooms assigned to the control group. During randomization, classrooms were blocked by program location (one of two locations) and by center. Children were then selected within each classroom. If we received consent forms for more than 10 eligible children within a classroom, the study team randomly selected 10 to be in the main sample.

Data Collection and Analysis (Efficacy Study):
Measures. Children were assessed one-on-one before and after the implementation of the *FSL*

professional development course, in the fall (Oct.-Nov.) and the spring (Apr.-Jun.). In this paper, we discuss findings from prediction trials on the *Preschool Assessment of Science* (*PAS*), a measure of preschoolers' knowledge and inquiry skills in physical science. Prediction tasks measure children's ability to accurately predict a physical event based on critical observation of related events. Internal consistency using Cronbach's alpha (α) is 0.727.

Analyses. Logistic regression analyses were conducted in order to determine the effect of *FSL* on specific items in the "Floating & Sinking", "Marbles & Ramps", and "Water Level" prediction tasks. Prediction tasks included three basic items: children's predictions of a physical phenomenon before it occurs (initial prediction), their ability to accurately observe that physical phenomenon as it occurs (observation), and finally their ability to revise or confirm their initial prediction based on their observation in order to make a related prediction (revised prediction). Because these items are binomial (either correct or incorrect), logistic regression was used to estimate the change in probability of correct response associated with predictor variables.

Findings / Implications:

Findings. In order to evaluate the effect of the curriculum on children's responses, children's responses in the spring were treated as outcomes, while controlling for their responses in the fall. Models were built incrementally with blocks of related variables, beginning with child-level covariates that were hypothesized to contribute to performance (age, gender, language status, IEP status, maternal education, fall score on outcome variable, and cognitive flexibility), followed by the classroom-level variable of intervention condition, and finally, interaction terms between participation in *FSL* and each child-level covariate. Individual predictors and interaction terms were evaluated using the Wald Chi-square statistic of regression coefficients (Peng 2002). Interaction terms were excluded from the final model if the Wald test was non-significant at p of .10 or more. Because the study design was nested, with children nested in classrooms, standard logistic regression may underestimate standard errors and overestimate effects. To account for this possibility, hierarchical linear modeling (HLM) with Bernoulli estimation for binomial outcomes was used to conduct confirmatory tests of final models obtained using logistic regression. Significant associations are presented for both logistic regression and HLM analyses.

In the "Floating & Sinking" task, children who participated in *FSL* were significantly more likely to make correct initial predictions, observations, and revised predictions. For the initial prediction responses, $B = 1.03(0.45)$, *Wald chi-square* = 5.34, $p < .05$, controlling for all covariates. There were significant negative interaction effects between participation in *FSL* and gender, $B = -1.01(0.50)$, *Wald chi-square* = 4.08, $p < .05$, and between participation in *FSL* and maternal education, $B = -1.46(0.64)$, *Wald chi-square* = 5.19, $p < .05$. These interactions indicated that girls and children with mid-level maternal education (high school diploma and/or some college) who participated in *FSL* had a greater increase in the likelihood of correct initial prediction from fall to spring compared to boys and children with low maternal education (less than a high school diploma) in *FSL*. These results were confirmed by HLM analyses. For the observation responses, $B = 0.52(0.25)$, *Wald chi-square* = 4.29, $p < .05$, controlling for all covariates, although the effect was very small. There were no significant interaction effects associated with participation in *FSL*. These results were confirmed by HLM analyses. The effect of *FSL* was only marginally significant according to HLM analyses, $B = 0.51(0.31)$, $t(61) = 1.67$, $p = .10$. For the revised predictions, $B = 0.81(0.32)$, *Wald chi-square* = 6.37, $p < .05$, controlling for all covariates. There were no significant interaction effects associated with participation in *FSL*. These results were confirmed by HLM analyses.

In contrast, we found mixed results on the other two tasks. In the "Water Level" task, children who participated in *FSL* were significantly more likely to make correct initial predictions and observations, but not revised predictions. For the initial predictions, B = 0.48(0.24), *Wald chi-square* = 3.94, $p < .05$, controlling for all covariates, although this effect was very small. There were no statistically significant interaction effects associated with participation in *FSL*. The effect of *FSL* was only marginally significant according to HLM analyses, B = 0.45(0.25), $t(61) = 1.82$, $p = .073$. For the observations, B = 1.01(0.37), *Wald chi-square* = 7.26, $p < .01$, controlling for all covariates. There were significant negative interaction effects between participation in *FSL* and low maternal education, B = -1.72(0.67), *Wald chi-square* = 6.59, $p < .05$, and between participation in *FSL* and high maternal education, B = -1.44(0.60), *Wald chi-square* = 5.90, $p < .05$. These interactions indicated that children with mid-level maternal education (high school diploma and/or some college) who participated in *FSL* had a greater increase in the likelihood of correct response from fall to spring compared to children with low maternal education (less than a high school diploma) or high maternal education (bachelor's degree) in the intervention. Finally, for the "Marbles & Ramps" task, we found no significant group differences on any of the items based on participation in *FSL*.

Implications. Although several items of this task showed evidence of an effect of *FSL,* several items did not. These conflicting results may be explained by 1) a lack of sensitivity of the measure to detect variation in children's knowledge and inquiry skills and/or 2) an insufficient strength of the intervention to change children's knowledge and inquiry skills. In order to ensure reliability of administration for the purpose of summative assessment, this measure was intentionally designed to constrain the manner in which children could respond (e.g. answering closed-ended questions, choosing one object among an array of objects). This assessment format, then, also constrains the degree to which children can demonstrate their learning, which may manifest itself through varied forms of expression (e.g. in conversation, in drawings, and in behavior). In contrast, instruction-embedded formative assessments, currently being piloted, will allow teachers to assess children's learning within the classroom context in a way that is both more authentic and more flexible than summative assessment techniques. In particular, through the use of "evidence forms," teachers may prompt children to make observations or predictions about a physical phenomenon (e.g. what water will look like in a tube when it is tilted; how blocks used for building structures may differ in shape or material). (Please insert Figure 1 here.) Children who are more verbal may respond by describing their prediction; others may draw a picture; still others may pantomime. By probing further, teachers can gain greater insight into children's learning by encouraging them, not only to demonstrate their knowledge but, to demonstrate their ability to answer questions through active manipulation of materials (e.g. putting water in the tube and seeing what happens).

Conclusions:

Although our results provide evidence for a positive impact on *FSL* on preschool children's understanding of science and scientific thinking, they are also mixed—a common situation in educational research. In this paper, we highlight a promising approach to learning from mixed (summative) results—by using formative assessment to provide more nuanced information about children's learning over time and to allow teachers to tailor instruction so that it is more effective.

Appendices

Not included in page count.

Appendix A. References

Arbuckle, J. L. (2010). *AMOS 19.0 User's guide*. Chicago, IL: Amos Development Corporation.

Ball, D. L. (2000). Bridging practices: Intertwining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education, 51*(3), 241-247.

Barnett, W. S. (2003). Better teachers, better preschools: Student achievement linked to teacher qualifications. *Preschool Policy Matters*,1-12. Retrieved from http://nieer.org/resources/policybriefs/2.pdf

Bowman, B., Donovan, M. S., & Burns, M. S. (2001). *Eager to learn: Educating our preschoolers*. Washington, DC: National Academy Press.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). *Taking science to school: Learning and teaching science in grades K–8*. Washington, DC: The National Academies Press.

Dwyer, M. C., Chait, R., & McKee, P. (2000). *Building strong foundations for early learning: The U.S. Department of Education's guide to high-quality early childhood education programs*. Washington, DC: U.S. Department of Education, Planning and Evaluation Service.

Espinosa, L. M. (2002). High quality preschool: Why we need it and what it looks like. *Preschool Policy Matters*,1-10. Retrieved from http://nieer.org/resources/policybriefs/1.pdf

Helburn, S. W., & Bergmann, B. R. (2002). *America's child care problem: The way out*. New York, NY: Palgrave.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Strutural Equation Modeling, 6*(1), 1-55.

Hulleman, C. S., & Cordray, D. S. (2009). Methodological studies. Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness, 2*, 88-110.

Loucks-Horsley, S., Love, N., Stiles, K. E., Mundry, S., & Hewson, P. W. (2003). *Designing professional development for teachers of science and mathematics* (2nd ed.). Thousand Oaks, CA: Sage.

Myers, D., & Dynarski, M. (2003). *Random assignment in program evaluation and intervention research: questions and answers*. Retrieved from http://www.ed.gov/rschstat/eval/resources/randomqa.pdf

Penuel, W. R., & Gallagher, L. P. (2008). *Comparing three approaches to preparing teachers to teach for deep understanding in Earth science: Short-term impacts on teachers and teaching practice*. Menlo Park, CA: SRI International. Retrieved from http://ctl.sri.com/publications/downloads/TIDES_ComparingThreeApproaches_EarthScience.pdf

Puma, M., Bell, S., Shapiro, G., Broene, P., Cook, R., Friedman, J., & Heid, C. (2001). *Building futures: The Head Start Impact Study. Research design plan*. Retrieved from http://www.acf.hhs.gov/programs/opre/hs/impact_study/reports/impact_study/impactstdy_resrch_plan.pdf

Shulman, L. S. (1987). Knowledge and teaching: Foundations for the new reform. *Harvard Educational Review, 57*, 1-22.

Smith, M. W., & Dickinson, D. K. (1994). Describing oral language opportunities and environments in Head Start and other preschool classrooms. *Early Childhood Research Quarterly, 9*, 345-366.

The White House. (2009). *Educate to innovate*. Retrieved from http://www.whitehouse.gov/issues/education/educate-innovate

Whitebook, M. (2003). *Early education quality: Higher teacher qualifications for better learning environments - a review of the literature*. Berkeley, CA: University of California, Berkeley. Retrieved from http://iir.berkeley.edu/cscce/pdf/teacher.pdf

Appendix B. Tables and Figures

Not included in page count.

Figure 1. Example of Instruction-Embedded Formative Assessment (*Evidence Form for Open Exploration of Structures*)

Session 2: Open Exploration of Structures

Evidence Form

Date: ____________________________

Participating Children:

1) ________________________________

2) ________________________________

3) ________________________________

4) ________________________________

Concepts **Solids have specific physical characteristics in common with one another.** **Building materials have characteristics that affect how effectively they can be used.**	
Goals	Indicators
o **Explores varied building materials, experiencing their differing characteristics and finding out what they can do with them** o **Develop purposeful engagement with building materials from which interests and questions emerge** o **Build capacity to participate in conversations about materials and building, generating a shared vocabulary about structures**	• Uses a variety of blocks • Engages with blocks for extended periods • Attempts to solve problems that arise • Shares ideas about the materials and process of building • Uses some descriptive vocabulary for explaining block characteristics and comparing similarities and differences

Please record your observations of children's behavior and conversation. When appropriate, attach children's representations.

Child Behavior	***Child in Conversation***
❶**Prompt: *We're going to spend time building with blocks today. How would you like to use the blocks?***	
❸**Prompt: *Here are two different kinds of blocks*** (show blocks of different shapes and materials). ***How are they different?***	
❹ **Prompt*: Tell me about what you built today? Which blocks did you use?***	

❺ **Observe: Describe how children respond during block play, noting types of structures built and blocks used. Choose 2-3 children and time how long they are actively engaged in the block area.**	
Additional Prompts:	

Session 2: Open Exploration of Structures

Do you have evidence of meeting your goals?

Looking at the evidence you collected on the **Evidence Form**, indicate how well each *Learning & Inquiry Goal* was met, citing *Evidence Item* numbers (❶-❺) and *Evidence Formats: Child Behavior **(B)**, Child Representation **(R)*** or *Child in Conversation **(C)***. For *Learning & Inquiry Goals* marked ✓- or ✓, indicate *Next Steps* you will take to help children reach these goals.

Indicators	Examples	Evidence of Meeting Goals? (✓- , ✓, ✓+; Evidence Item #s & Formats)	Next Steps
• **Uses a variety of blocks**	Child builds a structure using both kapla blocks and unit blocks; Child uses different blocks than he/she had been using previous day		
• **Engages with blocks for extended periods**	Child is focused on blocks for 20-30 minutes; Child chooses to stay in block area for duration of choice time		
• **Attempts to solve problems that arise**	Child notices that a tower of soft blocks keeps falling over and tries to use other blocks		
• **Shares ideas about the materials and process of building**	Child converses with peer about where to place blocks as they build a structure Child responds to teacher question about why he/she chose a certain block		