Abstract Title Page

Title:
Measurement of Fidelity of Implementation to a Core Technology Component and Effects on Outcomes in an Early Mathematics Intervention

Authors:
Julie Sarama (PRESENTER and CONTACT) (jsarama@buffalo.edu), University at Buffalo, SUNY
Douglas H. Clements (PRESENTER) (clements@buffalo.edu), University at Buffalo, SUNY
Mary Elaine Spitler (mspitler@buffalo.edu), University at Buffalo, SUNY
Christopher B. Wolfe (cwolfe1ster@gmail.com), Indiana University - Kokomo
Alissa Lange (alissaanne@hotmail.com), National Institute for Early Education Research (NIEER)
Abstract Body

Background / Context:
This research project's theoretical framework, undergirding the scaling up of a research-based mathematics intervention, is an interconnected model of interacting influences, with faithful implementation at the center of the model, linked to child outcomes (Clements, Sarama, Spitler, & Wolfe, 2011). Thus, fidelity of implementation (FOI) is hierarchically included within this model of networked influences.

Measurement of FOI is important in distinguishing between theory failure and implementation failure (Raudenbush, 2007). The FOI construct has evolved from a notion of assumed fidelity (Dearing, 2008; Rogers, 2003) to one of interrelated components that includes the complexity of context, the variance of enactments, the role of valid measurements, and the link between FOI and student outcomes (Berman & McLaughlin, 1976; O’Donnell, 2008). We operationalized the definition of fidelity as a determination of how well an innovation is implemented according to its original program design (Lee, Penfield, & Maerten-Rivera, 2009).

Researchers continue to seek a standard model for reliable and valid measurement that can account for the complexity of FOI, the differences in contexts (Zvoch, 2009), and the mediating and moderating effects of student (Raudenbush, 2007) and teacher characteristics (Achinstein & Ogawa, 2006). Recognizing the philosophical and practical tensions between replication and adaptation, researchers have called for the separate measurement of the effects of fidelity to, and adaptation of, the components of the intervention (Morrison et al., 2009; O’Donnell, 2008). In a similar vein are efforts to identify critical components of interventions (Mills & Ragan, 2000; Morrison et al., 2009; O’Donnell, 2008), with the understanding that evaluation theory is guided by program theory (Donaldson & Lipsey, 2006), which guides the determination of core components (O’Donnell, 2008). These researchers argue that critical components should be hypothesized a priori, and the extent to which teachers are faithful to each specific core component should be measured separately (O’Donnell, 2008; Shapley et al., 2010).

Building on this body of work, we hypothesized that FOI to the technology component, a hypothesized critical component of this intervention, would predict child outcome scores. We examined teachers’ fidelity to the technological elements using six indices, using data collected via a curriculum fidelity classroom observation instrument, mentor reports, teacher self-reports, and three sets of automated, computer-generated fidelity data. We then applied hierarchical linear modeling to examine the impact of fidelity measures on multiple child-level outcomes.

Purpose / Research Questions
Our goal is to contribute to research on fidelity of implementation by providing an empirical example addressing the theoretical trend toward distinguishing the core components of fidelity of implementation a priori (O’Donnell, 2008), developing and utilizing multiple measures to measure the extent to which there was faithfulness to a specific core component, and linking the measure to outcomes. Our example concerns the technology components of a research-based prekindergarten mathematics intervention, and the extent of implementation fidelity to child outcomes on a mathematics assessment, in the context of a scaled-up study of educational effectiveness.

Setting:
The study took place in pre-K classrooms in two urban school districts, the Buffalo Public School system in Buffalo, NY and the Boston Public School system in Boston, MA.
Participants:
Teachers (106, 72 experimental) from 42 public schools in two cities (one distal) participated in this longitudinal prekindergarten mathematics curriculum and professional development scale-up study. Eligibility for schools included stable feeder patterns, and no former participation by teachers in research projects involving the intended intervention. Participants were four-year-olds (51% Female) of mixed ethnicity (53% African American, 21% Hispanic, 19% White, 3.7% Asian Pacific, 1.8% Native American, and .6% Other). Most (82.33%) received free or reduced lunch, 13.5% had limited English proficiency, and 10% had an IEP. Eligibility of children was linked to attendance at schools that were randomly assigned to clusters.

Intervention:
The instructional approach of our early mathematics curriculum, Building Blocks is finding the mathematics in, and developing mathematics from, children’s activity (Clements & Sarama, 2007, 2008). At the core of Building Blocks’ content and sequencing, and the core of the CRF, are empirically-established learning trajectories. We defined learning trajectories as “descriptions of children’s thinking and learning…and a related, conjectured route through a set of instructional tasks” (Clements & Sarama, 2004, p. 83)

A central component of the curriculum is the Building Blocks Software, also based on the CRF (Clements, 2007) and its core of learning trajectories, following a detailed set of procedures specifically created to support scientifically-based software development (Clements, Sarama, Yelland, & Glass, 2008). As an example, the a priori phase of research reviewed showed that computers can be used effectively by children as young as 3 or 4 years of age and that software can be made more motivating and educationally effective by, for example, using animation and children’s voices and giving simple, clear feedback. As another example, the design of learning trajectories for the software emphasized that each activity must engender those mental processes or actions hypothesized to move children to the next level of the trajectory. The resulting software included over 60 activities and a management system that presents tasks, contingent on success, along the research-based learning trajectories. Teachers were expected to electronically assign to children a certain number of software activities (usually two) per week, aligned with the activities in the print curriculum. Children were expected to rotate to the computer during free or small-group time, and were expected to complete two computer activities per week.

The intervention included professional development, both 14 school-day training sessions over two years, centered around the mathematics curriculum, the technology component, and content, and coaching. Pre and post questionnaires were used to measure teachers’ attitudes and experiences with mathematics. After the child pre-test, teachers implemented the 30-week mathematics curriculum, delivering four components: whole group, small group, hands-on activity centers, and the computer/technology component.

Research Design:
In a Cluster Randomized Trial design, schools were publicly assigned to one of three treatment groups (using a table of random numbers, with blind pointing to establish the starting number). Two of the treatment groups were identical at pre-K one included follow through at Kindergarten and first grade), resulting in 72 classes in the experimental group and 34 in the control group.

Previous reports revealed statistically and practically significant effects of the intervention in mathematics outcomes (effect size = .72, Clements, Sarama, Spitler, Lange, & Wolfe, 2011) and language (effect sizes ranging from .13 to .29) (Sarama, Clements, Lange, & Wolfe, 2010). The current analyses focus on data from treatment classrooms only, as those classrooms are the only ones that implemented the specific computer activity component. We hypothesized that fidelity to the technology component would
significantly impact mathematics, but not language scores (given that all expressive language was embedded in the non-technology components).

**Data Collection and Analysis:**

We collected data on FOI using several measures, all designed to provide indices of faithfulness to the intervention. We also collected child outcome data using a measure of mathematics achievement.

*Curriculum Fidelity Classroom Observation Instrument*

This 52-item instrument measured the extent to which teachers faithfully implemented the curriculum. It was organized into five sections (General Curriculum, Hands-On Center Activities, Whole Group Activities, Small Group Activities, and Computer Activities). The “Computer Activities” subscale was used in this analysis. This scale consisted of 13 Likert-style questions. Two examples are: ‘The teacher’s classroom management strategies enhanced the quality of the activity and children’s mathematical learning;’ ‘The teacher knows how to access computer records of individual children’s sessions that are stored on the computer.’ Project mentors collected data using this instrument, after participating in several training meetings and interrater reliability visits over two years.

*Mentor Reports on Software Usage*

Classroom mentors rated teachers on the extent of their software implementation. Technology mentors independently rated the same teachers on the same topic using the same scale. These scores were averaged to provide an overall index rating for teacher software use specifically related to the intervention.

*Computer –Based Data*

We used data from three separate reports for this analysis: total time children spent on computers during the school year. For ‘Total Time on Computers’ start and stop times per each activity assigned or attempted by the individual child was recorded. Teacher level computer data was gleaned from the number of teacher logins onto the software site over the course of the academic year, and the percentage of weeks, out of 30, electronically assigned to the class.

*Teacher Questionnaire*

Three items on the self-reported questionnaire were relevant to implementation of the technology component. 1. ‘I use on-line resources to find materials relevant to my curriculum’; 2. ‘I use technology to maintain student records’; 3. I use technology to monitor student performance.’ These items were analyzed both as a subset and individually.

*Child Outcome (Mathematics Assessment)*

The standardized instrument used was the Tools for Early Assessment of Mathematics (TEAM) which uses an individual interview format, with explicit protocol, coding, and scoring procedures. It assesses children’s thinking and learning along research-based developmental trajectories within areas of mathematics considered significant for preschoolers, as determined by a consensus of participants in a national conference on early childhood mathematics standards, rather than mirroring the curriculum’s objectives or activities. It was previously submitted to the Rasch model entering only correctness scores, yielding a reliability of .94 (Clements, Sarama, & Liu, 2008). Data were collected on children’s mathematics achievement via pretests and posttests during their prekindergarten year.

**Findings:**

This analysis of fidelity to the technology component of the intervention was conducted on treatment only schools, as this technology component was unavailable to the control schools. Using two-level
Efficacy and Effective Studies of the Building Blocks Curriculum

hierarchical linear modeling, we examined the impact of fidelity measures on child-level posttest scores, after controlling for pretest scores. The pretest was a significant predictor of posttest score.

**Curriculum Fidelity Technology-Subscale**

Results indicated that FOI to the technology component of a mathematics curriculum, as measured by faithfulness to a technology subscale, positively and significantly impacted child test scores. The curriculum fidelity technology-subscale was hypothesized to be the primary measure of fidelity to the curriculum in general, and to the technology curriculum specifically. At level 2, using classically arranged scores, the 13-item technology-related subscale school aggregate was positively and significantly related to posttest scores, controlling for pretest (coefficient 0.29, SE=0.10, \( p < .008 \)). There were no significant main effects for gender, although a trend was detected toward males overall scoring lower overall. Significant interactions with the curriculum fidelity subscale in the prediction of child posttest scores were found. There were significant interactions with African Americans, Hispanics and Whites such that greater fidelity was related to higher post-test outcomes for these groups. A similar interaction was also found for children’s “Total Time on Computer.” These interactions will be further explicated in the presentation.

**Total Time on Computers**

Child curriculum software usage at Level 1 was also examined as a predictor of posttest outcomes. At level 1, children’s ‘Total Time on Computers’ was found to be a statistically significant, though non-robust, predictor of child mathematics posttest scores. Neither school-level SES nor child-level IEP differed significantly for this variable. In the prediction of posttest scores, controlling for pretest scores (treatment group only), there was not a significant interaction between site and ‘Total Time on Computer.’

**Other Measures**

We analyzed FOI to the technology component with five other measures, aggregated to the school level. Teacher Assignments of the weeks of the curriculum, computed by the percentage of weeks assigned, did not significantly impact on children’s posttest scores over and above the technology sub-scale, nor did mentor ratings on teacher software usage. The third measure, teacher logins, was found to have a small, negative impact. Neither the individual questions nor subset of the teacher self-reports had a significant impact on posttest scores, over and above the technology-subscale measure.

**Language Outcomes**

As hypothesized, none of the measures of FOI significantly impacted language scores.

**Conclusions:**

The main findings from this study support the hypothesis that fidelity of implementation to the technology component of a mathematics curriculum, as measured by faithfulness to a technology subscale, positively and significantly impacts child test scores, confirming other fidelity studies. This finding affirms the findings of Ysseldyke et al. (2003) in terms of the impact of faithful implementation of technology components, and its effect on children’s mathematics achievement. It also affirms technology as a core component of the intervention. Finally, findings suggest that observational measures and software-collected total time using the software may be more useful measures of FOI than mentor reports, teacher logins, or percentage of software modules assigned by the teacher.
Appendix: References


Efficacy and Effective Studies of the Building Blocks Curriculum


