

2009 CCPRC Annual Meeting
Breakout C-3
Friday October 30, 2009, 10:30 a.m. – 12:00 p.m.

Measuring Implementation in Research and Practice

Description

The session focused on measurement and evaluation of implementation, an essential factor in evaluating child care quality and its effects on children, parents, and families. Specifically, it addressed measurement and evaluation of fidelity—determining whether an initiative or intervention as a whole, the staffing and the services they offer, or a particular component such as a curriculum, is delivered in a way that is “faithful” to the initial intent and its anticipated outcomes.

This is an important issue for child care researchers and policy makers because replication of a model--or taking it to scale—depends on true understanding of its results. Researchers need to know the kinds of data that are important to collect and how to collect them; policy makers need this kind of information for decision-making about future directions. And, as a field, we have not paid a great deal of attention to this issue.

Specifically, this session provided a conceptual model for the measurement of fidelity of implementation of quality interventions. Then, presentations addressed quality interventions at two levels: 1) professional development targeted at improving teaching, and 2) curriculum models targeted at improving learning. Highlighting findings from evaluations of child care quality that examined fidelity in implementation, this session raised questions designed to inform future agendas for child care research.

Moderator:

Kate Tarrant, Columbia University

Panelists:

Sandra Naoom, National Implementation Research Network, Frank Porter Graham Child Development Institute, University of North Carolina at Chapel Hill

Ann Collins, Abt Associates, Inc.

Carolyn Layzer, Abt Associates, Inc.

Scribe:

Barbara Coccodrilli Carlson, National Child Care Information Center

1. Documents in Session Folder

- “Assessing Fidelity: A Necessary Component of Effective Implementation,” Sandra F. Naoom.
- “Massachusetts Family Child Care Study: Measuring Fidelity of Implementation,” Ann Collins.
- “Measuring Implementation in Research and Practice: Developing an Implementation Measure for Project Upgrade,” Carolyn Layzer.

2. Summary of Presentations

- **Introduction: Kate Tarrant**

- This presentation addressed the measurement and evaluation of “fidelity,” determining whether an initiative or an intervention or the staffing and services it offers (e.g., a particular component, such as a curriculum) are delivered in a way that is “faithful” to the initial intent and its anticipated outcomes.

- **Summary of Presentation #1: Sandra Naoom**

- *Stages of Implementation*: Exploration, installation, initial implementation, full implementation (these first four take 2-4 years), innovation, and sustainability.
- Often summative evaluations of a program’s worth or effectiveness occur before the program has been fully implemented, and is not a good indicator of the potential of the program.
- *Implementation “Drivers/Processes”* include recruitment and selection, pre-service training, consultation and coaching, systems interventions, facilitative/administrative supports, etc. (see PPT for complete list).
- High-quality fidelity rarely occurs in “natural” settings; it’s more likely to occur in controlled research settings.
 - We measure fidelity to determine if there is an implementation or effectiveness problem.
- Dimensions of intervention fidelity: Adherence, dosage, quality of program delivery, participant responsiveness, and program differentiation.
- National Implementation Research Network’s dimensions of fidelity: Context, compliance, and competence.
- Fidelity assessments should address structure (the framework for service delivery) and process (the ways in which services are delivered).
- Methods to develop a fidelity assessment: Must be able to specify the Theory of Change or the core components of intervention need to be identified and measured.
- Common methods of measuring fidelity: Observation, which is the gold standard; ratings by experts, and consumer self-reports.

- **Summary of Presentation #2: Ann Collins**

- Looked at fidelity of Learning Games curriculum implementation with family child care providers in Massachusetts.
- Study involved 350 family child care network providers.
- 50% of these providers were randomly selected to use the Learning Games curriculum.
- Study measured implementation through observation, the TALK measure (created for study), the Arnett, and nine additional items from other measures.
- Implementation involved two components: Professional Development (PD) and Program Model.
 - Program model: Created an implementation scale (see PPT slide #20).
 - Fidelity study found that the PD model was only partially implemented.
 - Provider observations demonstrated a slightly higher score on the implementation scale.

- **Summary of Presentation #3: Carolyn Layzer**
 - Focused on implementation of Project Upgrade: A test of three language and literacy curricula in Miami/Dade, Florida.
 - It would be great if the designer of an intervention would include a tool to measure fidelity.
 - Included vertical (going in depth into a particular curriculum) versus horizontal (used across curricula and conditions) measures of implementation.
 - Each developer in the Project Upgrade curricula used a fidelity measurement specific to his or her intervention.
 - Abt used measures from the OMLIT battery, including OMLIT-Snapshot (A Snapshot of Classroom Activities), to measure proximal outcomes.
 - When coaches and mentors (use developer tools to) rate the teachers with whom they work, they tend to give higher ratings in general; they have difficulty being objective.
 - See PPT slides 6–10 for two of the three curricula; the process Abt went through is described to identify variables to include in implementation measures and instruments. Broadly, the process was to identify OMLIT variables that matched the aspects of the curricula listed in the developers’ own rating tools to create an index describing the level of implementation observed by independent observers.
 - Caveats: Important aspects of implementation, such as intensity, exposure, or quality of interaction, are not always explicit in a developer’s measure and are also difficult for objective observers to measure.

3. Summary of Discussion

- The replication of a model and the process of taking it “to scale” depends on a true understanding of the results. This approach is most important when people don’t find an overall impact, or the effect is small. Some people do it well, some less well. This is what happens in the real world. Turnover is a *big* issue, even in the control classrooms; we can’t lock teachers into place.
- The line between fidelity, level of implementation, and good practice depends on the intervention used. This work is very important: A professional development project in Minnesota found it difficult to track how well the providers were doing. We need to have a better understanding of when we get an “outcome” and when we don’t. We rarely have a record of the practices that are linked to implementation.
- A question was asked about what venues can be used to make the field aware of the importance of this kind of fidelity work. Responses included:
 - We need to plan more at the outset of research projects to ensure that outcomes will be of interest.
 - This is feasible when looking at a single intervention. Imagine designing a fidelity measure across three different interventions and an outcome measure as well. Even with infinite resources, this would be difficult to achieve.
 - There is tension between fidelity and adaptation to local circumstances. For example, in the work in Miami, teachers needed more mentoring than developers expected.
- Learning Games came up with a “fidelity” score, but it only represents one point in time.
- Most people look at an implementation study as an “add-on.” Fidelity and implementation should be thought about at the outset. We need to ask the questions:

- Who is responsible for what?
 - Is the evaluator responsible for documenting anything relevant?
 - How much money do you have to put the implementation piece in the overall evaluation?
- When do you administer a fidelity instrument?
 - This has not been figured out yet. Ongoing measures of fidelity are crucial because they give us information on how to improve implementation.
 - When we did a fidelity study on a small scale pilot, we learned an enormous amount, and went back to the drawing board for the larger study.

4. Key Themes and Issues

- These interventions are not happening in a vacuum.
- Our kids are dynamic and interventions change (e.g., dual language learners). We work in a dynamic context and we need to consider that.
- We need a venue for sharing best practices (e.g., measuring fidelity and identifying when intervention components are working).
- It's important to document contextual issues.
- Be thoughtful and plan fidelity measures when evaluating an approach to a large evaluation project.
- Feasibility varies depending on the intervention (e.g., technical issues and reliability across different settings).
- Tension exists between fidelity and adaptation.
- If you don't document implementation, it makes it difficult for others to replicate the intervention.
- Evaluating efficacy is important, but evaluating implementation may be more useful to the field.
- Fidelity measures allow us to learn lessons around real world implementation.
- CCPRC should think about short documents to advise policymakers on fidelity measures and their importance in "scale up" and "implementation" projects. Participants should think about working with NIRN and looking at their issue briefs.
- We need to look at implementing the model, tracking fidelity, then testing the outcomes for the intended beneficiary of the particular model.