



REFLECTIONS  
ON  
METHODOLOGY

DECEMBER  
2018

**NEW YORK**  
200 Vesey Street, 23rd Fl.  
New York, NY 10281  
Tel: 212 532 3200

**OAKLAND**  
475 14th Street  
Suite 750  
Oakland, CA 94612  
Tel: 510 663 6372

**WASHINGTON, DC**  
1990 M Street, NW  
Suite 340  
Washington, DC 20036

**LOS ANGELES**  
11965 Venice Boulevard  
Suite 402  
Los Angeles, CA 90066

www.mdrc.org



# Working with Practitioners to Develop Measures of Implementation Fidelity

By Meghan McCormick and Michelle Maier

*This post is one in a series highlighting MDRC's methodological work. Contributors discuss the refinement and practical use of research methods being employed across our organization.*

Social and education policy researchers often work in partnership with the practitioners who develop and implement the programs being studied. Engagement with those who design and use a program is particularly helpful in developing measures of implementation fidelity — that is, whether a program is being implemented as designed. There are two goals here: Researchers want to understand why an intervention did or did not produce effects, and practitioners want to use the information to make improvements beyond the study period.

One way researchers measure fidelity of implementation is through observation tools: for example, a protocol or rating form that allows a classroom observer to record how often a teacher employs a particular practice or method and assess the quality of that practice. But valid and reliable observation tools are often not readily available for practitioners' later use. By working together to construct the tools, researchers and practitioners both benefit.

## CONSTRUCTING THE EVALUATION TOOLS

MDRC has begun working with research partners through the [Expanding Children's Early Learning \(ExCEL\) Network](#), a collaboration of researchers, preschool providers, and local stakeholders, to construct tools to assess implementation fidelity in a set of early childhood education programs developed by Boston Public Schools. To date, this work has been conducted in three stages. In stage one, the research team reviews the components of a program or curriculum and, following a framework summarized by [Fixsen and colleagues \(2005\)](#), identifies a core set of practices or activities to measure within three domains:

1. **Dosage**, or the amount of time providers spend implementing the core components of the program, compared with the time recommended in the program model
2. **Adherence**, or the percentage of program model components implemented as intended (the extent to which the model is implemented)
3. **Quality**, or the extent to which the program model components are implemented using strong practices (the extent to which the model is implemented *well*)

In the second stage, the partners engage in an iterative process, making several rounds of edits until the researchers and the model developers agree that the implementation can be measured reliably and validly and that the data will be both an accurate assessment of fidelity to the specific model and useful to the practitioner. Part of this iterative process includes at least one trial observation in the field, after which the participating researchers and practitioners discuss the experience. The insights gleaned from this type of live observation have proven invaluable for further adapting the fidelity measures. Eventually, the partners agree on a new version of the measure.

In the third stage, the research team trains supervisors or coaches from the model developer to collect these data and code them properly. Those staff members may have further insights that

inform a final set of edits to the tool. The practitioners then engage in a set of activities to ensure that they can collect the data reliably in live observations. To test reliability, for example, researchers might ask staff members to code a videotaped observation and then compare the results with those previously obtained by a master coder. In the course of the evaluation, researchers and practitioners must come together often to discuss findings, to ensure that the data collected in the field are accurate, objective, and sufficiently rigorous to be used for research purposes.

## **BENEFITS AND CHALLENGES OF THE PARTNERSHIP**

There are several benefits to developing implementation fidelity tools in direct partnership with practitioners — besides the accuracy and utility discussed above:

- The tool will use the same terminology that staff members do.
- Large-scale data collection that may benefit researchers and practitioners alike can be conducted by staff members who know the programs well rather than by outside research teams.
- Practitioners' role in constructing the tool gives them ownership of it, which may make them more likely to integrate it into their practice on a regular basis, potentially improving program operations over time.

At the same time, there are challenges for both researchers and practitioners to consider:

- For fidelity tools to be relevant and easy to use, the protocols and rating systems themselves need to be concise and focused on the core practices that define the program model. Researchers and practitioners may have competing priorities for the tools, which may cause them to become too long and detailed.
- For the data collected with the observation tool to be useful and directly inform practice, they need to be analyzed soon after being collected, ideally within two or three weeks. Data processing, analyzing, and summarizing can use a lot of computational and staff resources. It is important to think about the implications for analysis when developing the tool.
- Programs evolve — sometimes on the basis of fidelity research — and model components change. Researchers and practitioners must recognize that fidelity tools will require further adaptation. Maintaining consistent training and reliability standards over the years can help address this challenge.
- Even with processes to support systematic and high-quality data collection, there will be variation in what is observed, on which day, how often, and for how long, making it difficult to compare observations. Teams can address this issue by developing tools that assess the core activities that should occur every day, or in any observation period. Such global measures can help capture the fundamental practices that are critical to the program model.

If practitioners can collect reliable and valid implementation data, those data can inform the program's daily decision making. At the same time, the data can help researchers learn more about the core practices that are hypothesized to explain positive outcomes for program participants.

In the [ExCEL project](#), MDRC is using implementation fidelity data for a preschool program to help Boston Public Schools understand how variation in fidelity affects children's academic gains over the course of a school year. That project will be discussed further in a follow-up post.