



outlier

RESEARCH & EVALUATION
CEMSE | UNIVERSITY OF CHICAGO

[SEE PRESENTATION NOTES AT END OF
THIS DOCUMENT]



THE UNIVERSITY OF CHICAGO

10

**Top Ten Considerations
for Evaluating Implementation Measurement
and Analysis Strategies**

5

Parts to Consider in the Implementation Process:

Theory/Background

Design

Measurement

Data Collection

Analysis

2

Principles to Keep in Mind: Rigor and Fit

Taking Care of Business



Intervention

Implementation Research

Intervention Implementation

Fidelity of Implementation

Implementation Factors

The Implementation Process

10

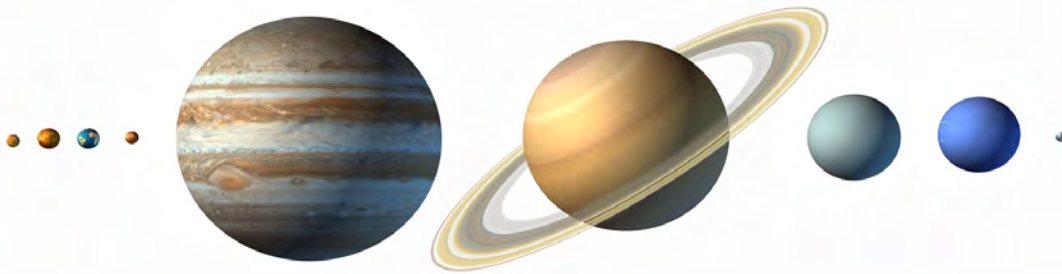
Theory and Background:

Is it there? Serious attention to Implementation measurement in education is relatively new.



9

Theory and Background:
Watch for clear, well-defined terminology.





8

Design:

Distinguish between intervention implementation and factors that affect intervention implementation.

7

Design: Define and measure the “it.”



6

Measurement:

Remember that interventions are multi-dimensional.





5

**Measurement:
Adaptation is a given (with some exceptions)
and isn't inherently bad.**



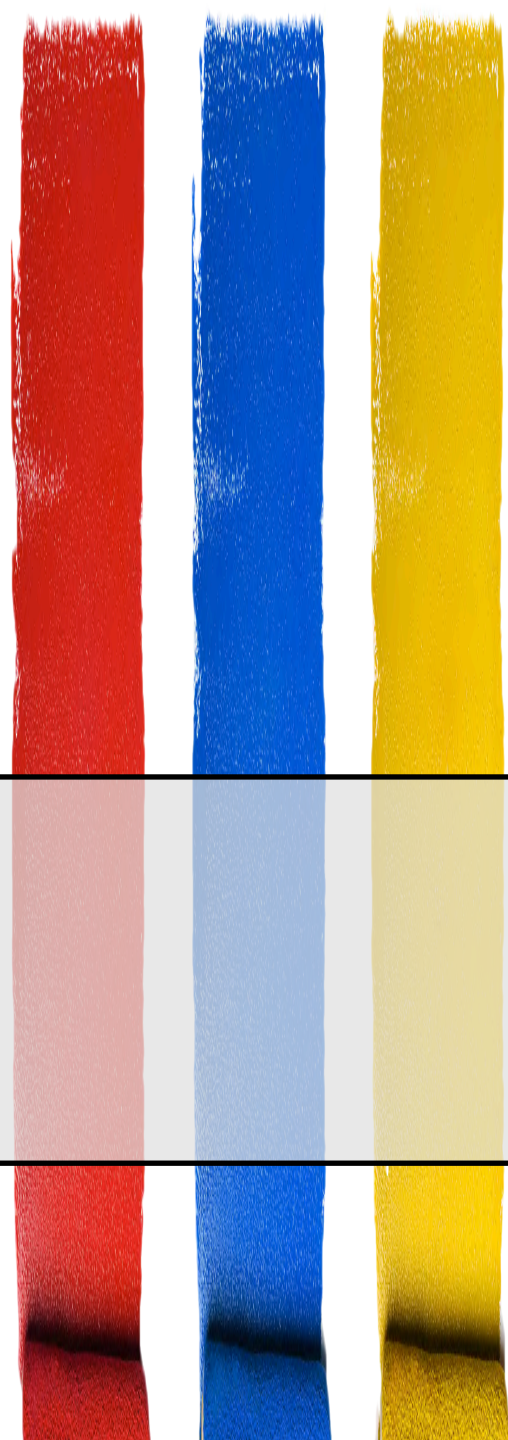
4

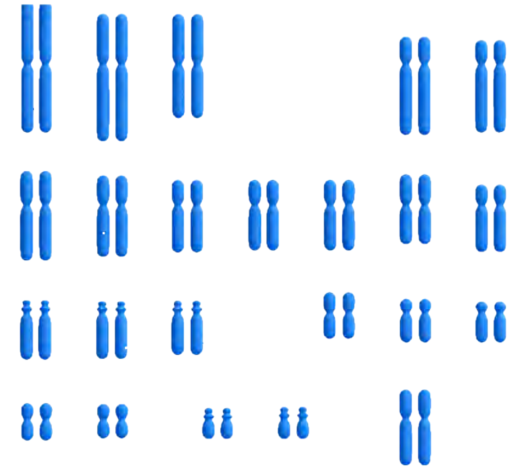
**Data Collection:
Be realistic and pragmatic.**



3

**Analysis (and Design and Measurement):
Use the right grain size.**





2

**Analysis:
Reduce responsibly**

1

**Theory and Background:
Build on theory, understand fundamental concepts
and contribute to knowledge.**





outlier

RESEARCH & EVALUATION
CEMSE | UNIVERSITY OF CHICAGO

Slide 1: I'm Jeanne Century, Director of Outlier Research and Evaluation – the research group at the Center for Elementary Mathematics and Science at the University of Chicago. Outlier seeks to create tools and knowledge to support and empower those seeking to advance education.

Slide 2: If you had the chance to take a look at the readings you will have seen that implementation measurement and analysis in education is still just emerging as an area of research and focus. You will also have gotten a glimpse of the complicated issues that confront those of us who want to measure and analyze implementation. The handouts give you some background on the work I've done in this area – much of which has been supported by NSF. I only have a half hour today so I'm not going to review the theory, background, research designs or scholarship of that work. Instead, I'm just going to give you the “top ten” considerations for you to keep in mind as you evaluate the implementation measurement and analysis strategies in proposals and in grantees and others' work.

Slide 3: My colleague, Amy Cassata, along with Chris Hulleman at the University of Virginia and Lisa Sanetti at the University of Connecticut have recently put in a proposal to IES for a Training Grant on implementation measurement. In organizing that work, we organized the implementation measurement process into five parts: Theory and Background, Design, Measurement, Data Collection, and Analysis. These are typical areas of research of course, but efforts to measure and analyze innovation implementation don't often address these different parts.

Slide 4: There are two overarching principles to keep in mind as you review these considerations: rigor and fit. By “rigor” I mean the thoroughness and thoughtfulness of the project and the extent to which it reflects theory, current knowledge and practice and by “fit” I mean the extent to which the implementation measurement efforts (and the parts of it) are an appropriate match for the overall study and its parts.

Slide 5: Before we move ahead any further, we need to “take care of business.” TCB. And in this world, that business is language. The failure to, and importance of having and using clear, consistent language could fill a talk of its own. So we'll just focus on what we need for today. These are all on your handout.

- Intervention: A general term for any program, reform, or innovation intended to change and improve practice.
- Implementation Research: The systematics study of the use of interventions in particular settings in order to create generalizable knowledge about their effectiveness, and/or how they operate and spread in natural settings.
- Intervention Implementation: The status of intervention use at a particular time.
- Fidelity of Implementation: The difference between intervention implementation as it is enacted and a theoretical ideal.
- Implementation Factors: The contexts and conditions that affect intervention implementation.
- The Implementation Process: A “big picture” of implementation comprised of both intervention implementation and implementation factors.

Slide 6: Attention to implementation in education is relatively new. Last year, we looked at the What Works Clearinghouse (WWC) database of research articles. We reviewed 649 publications archived in WWC from 2009 to 2012 (including 540 intervention reports, 68 quick reviews, 14 practice guides, and 27 reference resources). Of these, only 23 (3%) provide detailed information on implementation measurement and/or analysis. One might assume that the WWC archive would provide fertile ground for locating research that has rigorously addressed implementation, but it's rare. Look for people who are giving attention to theory; to understanding what is known – work has been done in other fields – and direct those who don't know to those resources. They have to pay attention to the theory and background or they will needlessly and wastefully reinvent.

Slide 7: Recognize that implementation, dissemination, spread, scale-up, and diffusion aren't synonymous. We've already talked about how implementation and fidelity are different. Our WWC review identified many inconsistencies in the ways that the implementation process is conceptualized, measured, and analyzed. The lack of common theory and terminology, shared measurement processes, and analysis strategies present a challenge. This is part of rigor. Language is a first step.

Slide 8: The research design has to demonstrate clarity in the differences between the intervention (and measuring it) and the factors that affect intervention implementation. Some researchers mix them and we address that in our written pieces that you have. In fact, sometimes it's not so clear (we'll get to that in a minute). But when it comes to design – it's essential that the proposers have conceptual clarity or they won't be able to execute or analyze clearly. This is part of rigor and fit.

Slide 9: Let's talk about defining the "it." I don't know if some of you will be surprised or if some of you will be nodding your heads along with me – but having been a developer a researcher and a regular evaluator, creators of interventions don't always have a clear picture of their model or theory of action. This has to be clear in order to know what the "it" is. It's not so easy. Sometimes, elements of implementation we might think are factors – such as professional development – are actually part of the model. We'll see that in one of our examples. Some of you know Larry Malone from the Lawrence Hall of Science. When we were first developing this work, we were looking at the FOSS program. I kept asking him – Larry – if they don't do this...is it FOSS? Or, if they don't do this...is it FOSS? He finally threw up his arms and said, "If it's good, it's FOSS!".

Slide 10: One of the fundamental concepts to consider is that interventions aren't "wholes" that are or aren't implemented...or even "wholes" that are partially implemented. Some literature focusing on implementation and diffusion of innovations, while very helpful, focuses on relatively simple innovations – such as using a stove. In social settings like education, it's important to recognize that interventions are comprised of components. And those components can be organized into dimensions. Some people refer to these components as "Critical Components" but we've moved away from that language because it actually has two meanings – those components that are determined to be critical by the creator – theoretically critical – and those that are empirically demonstrated to be critical – contribute to desired outcomes.

Slide 11: Now let's talk about adaptation. This is a big issue – not only for fidelity questions but for implementation measurement. We say that adaptation is a given – in fact, we take a position that it is essential for enduring change. If there is no adaptation, interventions/innovations do not last. But again, that is another topic. There are exceptions, of course – in very carefully controlled settings. But for the most part – it's always there. The challenge is giving attention to what those adaptations are. Once you consider that an innovation is comprised of components that reside in categories or dimensions, one can then examine whether the adaptation are helpful. There is the idea of principled adaptation – adaptations that preserve the theory of action; as opposed to "fatal adaptations" that go against the theory of action and impact effectiveness. In this area, it is also important to consider the difference between examining fidelity and examining use – adaptations from a desire for fidelity perspective are "bad" – from a "use" perspective, adaptation isn't inherently bad or good.

Slide 12: Data collection – be realistic and pragmatic. A district cannot force a school to participate and, if we are following IRB protocol, a school and any teachers within a school cannot be compelled to participate in any study. We are well aware of all of the strategies one can use to collect sufficient data – and it remains a challenge. Don't ask yourself “why don't they” or “why won't they..” ask yourself “why should they participate?” or “Why would they participate?” Ask yourself – if you were asked to do this and you didn't have to...would you? Why?

Slide 13. Now we get into analysis questions – the issue of grain size also relates to design and measurement but is very clear in the context of analysis. Grain size pertains not only to data collection (teacher, classroom, school, district) but also the grain size of the intervention (lesson, unit, curriculum) etc. Remember, that interventions are multi-dimensional and you won't be measuring all components with every instrument. Which components will you measure with what instruments? From whom will you collect data? How often? And this is closely tied to consideration number 2....

Slide 14: This slide is talking about reducing and aggregating data. It has been said that classroom observations are the gold standard of measuring implementation – but there are important elements of implementation that can't be measured in a single observation or two or three or four; depending on when those observations are made. We study reform-based curricula and these interventions have model components that aren't expected to be part of every lesson.

Researchers may observe vertically – one or two teachers in a school across grade levels...but what can that data tell you about implementation in the school? Or, researchers may observe horizontally – 2, 3, 4, or more observations of a single teacher. But what does that tell you about the teacher practice? How do you reduce the data?

Slide 15: Finally...we get back to theory. Meaning, researchers engaged with implementation measurement and analysis need to communicate to others what they are doing and engage in an on-going dialogue with others in the field. This is an opportunity to thoughtfully build knowledge – to not reinvent and a necessary part of a cycle of continuous growth. Other fields, other researchers know a lot and this is the time to reach beyond education and build expertise together.

Slide 16: I've given you a lot to consider. As you look at the scenarios in your handout – whether you want to join in the small group to discuss them or not – the most important thing to remember isn't the ten considerations or the five parts of the implementation measurement process – but the two principles – rigor and fit – and keep those in mind to frame your implementation measurement and analysis evaluation.