



**THE FUTURE OF
IMPLEMENTATION EVALUATION:
TOWARDS A SCIENCE OF
“CONTINUOUS IMPROVEMENT
RESEARCH IN EDUCATION”**

Barbara Schneider

John A. Hannah University Distinguished Professor, MSU
Senior Fellow, NORC

Purpose

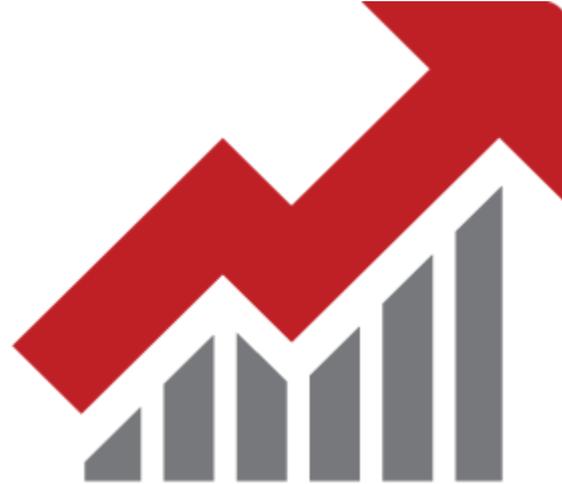
The primary objective of this day-long workshop at NSF is to:

- **Review** *rigorous methods* for program and project evaluation
- **Discuss** *challenges* of implementation evaluation in an impact framework

Workshop Structure

(1) Measuring impact

- Brief overview
- Causal analysis
- Understanding the measures
 - *Program effect*
 - *Proposal effect*



(2) Informing NSF's review of proposals

- Taking projects and interventions “to scale”
- Considering obstacles and issues
- Developing and advancing impact measures

Presentation Structure

- **Changes** in impact evaluation and study
 - A brief history
 - Shifting *focus* to impact
- **AERA and ARC's role** in measuring impact
- **IES/NSF guidelines** for evaluation
- Some **implications for scale-up**
 - Possibilities
 - Challenges

Changes in Evaluation and Impact Studies

- **Russ Whitehurst**

- Workshops at OMB
- Offering new perspectives on impact and policy evaluation as an important resource



- **Causal analysis of impact**

- Impact is now the *focus* of evaluation
- Readily used in local and regional-level intervention studies

Measuring Impact

- Methods for *rigorous impact evaluation* are becoming more **defined** and **sophisticated** in their approach
- With respect to **definition**:
 - *AERA* empirical guidelines for reporting
 - *ARC's* methodological standards
 - *IES* and *NSF* guidelines for proposals

ARC

- **Criteria and guidelines for rating methodological rigor**
 - *Contribution* to knowledge
 - *Design*
 - *Sources* of data
 - *Measures* and *classification* schemes
 - *Analyses* and *interpretations*
 - *Generalization*

IES/NSF Guidelines

Six Types of Research

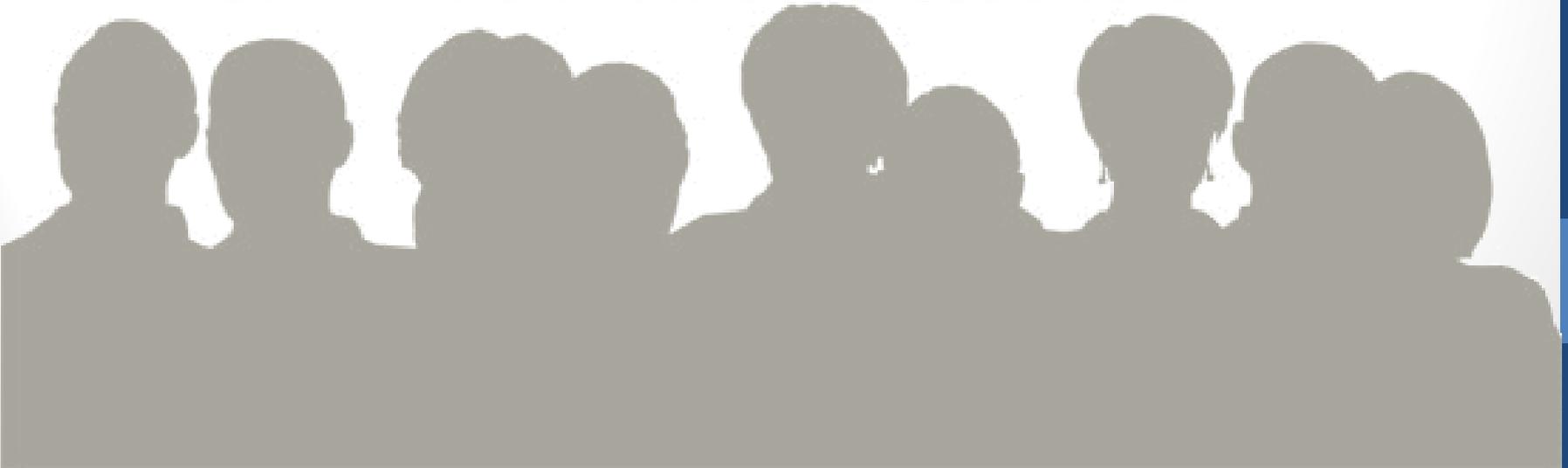
- **1. Foundation Research** –Studies of this type seek to test, develop, or refine theories of teaching or learning and may develop innovations in methodologies and/or technologies
- **2. Early Stage or Exploratory Research**-examines relationships among important constructs in education and learning to establish logical connections. These are usually correlational not causal.
- **3. Design and Development Research**-develops solutions to achieve a goal related to education or learning such as improving student engagement. Pilot tests of fully developed interventions warrant more advanced testing.

IES/NSF

- Purpose of Studies that Assess the Impact of Education Interventions and Strategies
 - **Efficacy Research**
 - Determine whether an intervention or strategy can improve outcomes under what are sometimes called “ideal” conditions
 - **Effectiveness Research**
 - Estimate the impacts of an intervention or strategy when implemented under conditions of routine practice
 - **Scale-up Research**
 - Estimate the impacts of an intervention or strategy under conditions of routine practice and across a broad spectrum of populations and settings

Going “To Scale”

- **Consider** the following:
 - *Almost 100,000* operating elementary and secondary schools
 - *50 million* students
 - *3.1 million* full-time teachers



Going “To Scale”

(continued)

- Focus on **considering the implications of going “to scale”** with programs and proposals
 - Decisions need to be made *earlier* in the research process
 - Creating *reasonable* expectations

Some Basic Considerations

- 1. Knowledge is not linear. There are important feedback loops that can inform and provide justification for other types of research.
- 2. Investigations can sometimes move directly from development of core knowledge to Scale-up.
- 3. Individual studies may incorporate elements that cut across research types. An efficacy study may need to engage in design and development cycles.

IES/NSF

- Guidelines for Evidence to Be Produced by Studies
 - **Project Outcomes**
 - Detailed descriptions of the study goals, design and implementation, data collection and quality, and analysis and findings
 - **Research Plan**
 - Study design used to estimate causal impact of the intervention on the outcomes of interest
 - Key outcomes of interest for the impact study and the minimum impact that would have policy or practical relevance
 - Study setting(s) and target population(s)
 - The sample
 - The data collection plan
 - The analysis and reporting plan

Recent Considerations

- Techniques are being developed to investigate different possible sources of variability found in multiple areas of research on the effects of programs.
 1. Variation imposed by researchers themselves
 2. Variation in implementation, settings, participants, and other features
 3. Variation unexplained or unaccounted for
- See **“A Conceptual Framework for Studying the Sources of Variation in Program Effects”** (Weiss, Bloom & Brock 2013)