What is Evidence?

Perspectives from the REESE program evaluation

Joy Frechtling, Westat

ITEST Summit 2012

“Evidence” –different frames at different levels

- Program--investment strategy for advancing understanding across a field or system
  - an accumulation, across projects, across themes/topics
  - broad, tied to full range of hoped for goals
  - extended in time, multiple funding cycles
  - context a subquestion
- Project--a targeted activity aimed at exploring an hypothesis or testing out a model/tool
  - an event
  - specific, tied to selected aspect of program goals
  - time limited
  - context specific?

- Very, very different....
- BUT...let's look more deeply at the REESE program evaluation and see how different they really are

Goals of the REESE program

- The goals of the REESE program are:
  1. *to catalyze discovery and innovation at the frontiers of STEM learning, education, and evaluation;*
  2. *to stimulate the field to produce high quality and robust research results through the progress of theory, method, and human resources; and*
  3. *to coordinate and transform advances in education, learning research, and evaluation.*

Evaluation Questions from NSF

- To what extent is the REESE program catalyzing discovery and innovation in research in STEM learning, education, and evaluation?
- How effective is the REESE program in stimulating STEM fields to produce high-quality, rigorous research studies (with rigor being appropriately defined to reflect the phase of the Cycle of Research and Development that they address)?

Evaluation Questions from NSF (continued)

- How effective is the REESE program in supporting the accumulation of knowledge in STEM education at all levels and in all settings?
- How effective is the REESE program in promoting interdisciplinary research in pursuit of national questions of STEM teaching and learning?

Challenges

Varied status in Cycle of R & D

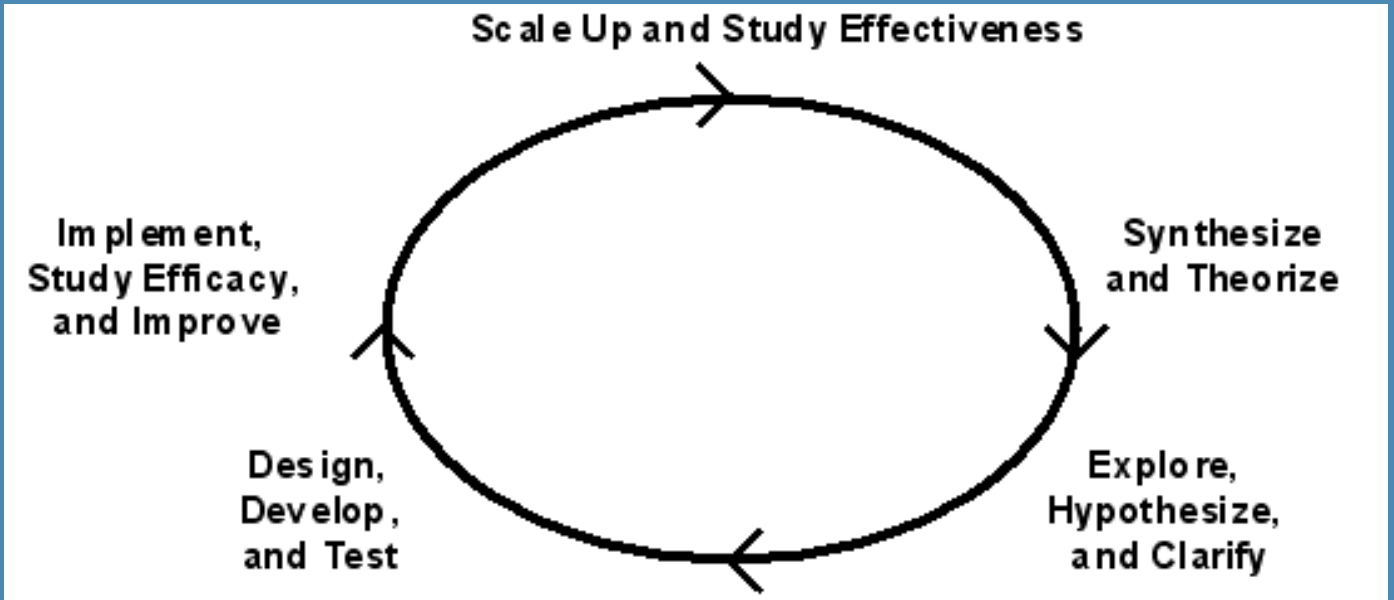

Figure 1: Cycle of Research and Development

Challenges (continued)

- Short time since program inception
- Attribution of outcomes to REESE
  - Separating impacts of REESE funding from those of other funding sources
  - Lack of a comparison group, benchmarks in some areas

Building the Evidence Base in Program Evaluation

- Recognize that most issues/problems need to addressed through multiple perspectives—multiple stakeholders, multiple interests
- Use multiple methods to explore hypotheses; triangulate sources of evidence; combine the traditional with the emergent

Building the Evidence Base

- Select designs/methodologies that meet standards of rigor and quality for type/focus of investigation
- Connect work to critical “Big Ideas” in the field; create the case for generalizability

For Example—for Question 1

- Is the REESE program identifying and funding research believed by the field to be important with regard to discovery and innovation in STEM teaching and learning?

Multiple Perspectives

- Is the REESE program identifying and funding research believed by the field to be important with regard to discovery and innovation in STEM teaching and learning?
- To what extent and in what ways is REESE impacting the focus of research proposals submitted to NSF to study STEM learning, education, and evaluation?
- Are projects that receive REESE funding able to use their REESE-supported work to obtain additional funding from NSF or other funding sources?
- Is REESE funding engaging new investigators and future investigators in STEM research on learning, education, and evaluation?

Multiple Methodologies

- Delphi panel—what are the hot issues?
- Review of solicitations—what is REESE asking for?
- Review of portfolio of funded projects—what is REESE funding?
- Review of a sample of unfunded projects—what is being proposed more broadly?
- Survey of PIs (funded and unfunded proposals)—how is the REESE project situated in researchers' broader research program

Quality and Rigor

- Fuzzy RDD
- Benchmarks from bibliometric data bases
- Use of previous established measures and instruments—review of research rigor

Big Ideas

- Need for interdisciplinary/team approach to address complex problems
- Dissemination of knowledge and building the knowledge base

What makes for evidence at the project level?

- Same principles apply, although the focus is different and the evaluation will look very different
  - Recognize that most issues/problems need to addressed through multiple perspectives—multiple stakeholders, multiple interests
  - Use multiple methods to explore hypotheses; triangulate sources of evidence; combine the traditional with the emergent
  - Select designs/methodologies that meet standards of rigor and quality for type/focus of investigation
  - Connect work to critical “Big Ideas” in the field; create the case for generalizability

SRI International

What is (your best) evidence?

Taking an NSF program perspective

Patrick Shields, ITEST Program Evaluation
Vera Michalchik, ISE Program Evaluation

NSF requires *real* evidence of *real* learning

- Science content based on standards
- Control conditions
- Random assignment

NSF wants knowledge about

- How to best support strong
  - Programming models
  - Learning outcomes

NSF wants STEM content outcomes

- Importance of standards over the years
- New standards highlight scientific process
- Science is what NSF is about

NSF has broadened view of outcomes

- Friedman *Framework* (2008)
  - Awareness/knowledge
  - Engagement/interest
  - Attititudes
  - Behavior
  - Skills

NSF has broadened view of outcomes

- *Learning Science in Informal Environments* (2009), expands on *Taking Science to School* to make:
- "Interest" foundational
- "Identity" central

Science comes first

- "Soft" outcomes have to be secondary
- We know how to test for science learning

Whole experience comes first

- Cognitive/affective don't split well
  - From each other
  - From context
- Need positive scientific experiences

Need to worry about transfer

- Out-of-school must transfer to school
- School makes science systematic
- School achievement is what counts

Need to worry about interest

- Motivation, engagement, persistence
- School take over afterschool?
- Afterschool influence school?

Need standardization of measures

- Avoid reinventing the wheel
- Demonstrate outcomes that matter
- Aggregate across programs

Need programmatic diversity

- Standard measures => standard programs
- Standard programs=> standard thinking

Need comparison groups

- Can’t do research without it
- Requires some degree of control

Need to compare what's there

- Take advantage of natural variation
- Use design based research approach
- Measures aligned with programming

So let’s treat and test

- Align measures with programming
- Administer tests

Let's not wreck the experience

- Norms for OST are different from school
- Different types of evaluative practices
- Programming comes first

Gotta break eggs to make omelet!!

- Everyone's used to taking tests
- We have good valid outcome measures
- We won't make progress your way

Gotta have compatible ingredients

- Validity is measuring what you mean to measure
  - Attraction, engagement?
  - Effects of the experience?
  - Capacity to
    - Use resources wisely?
    - Solve problems?
    - Take up future opportunities?

Must measure what's in the person

- Carries forward to new situations
- Let's us know if individuals are learning
- Let's us know if they'll do more next time

Measure uptake of opportunities

- Focus on choices, preparation, commitments, performance
- See these things in light of conditions
- How are we better preparing people to respond to their opportunities?

NSF MUST HAVE SUMMATIVE DATA!!

- Programs need it
- Congress needs it
- Taxpayers need it!!

NSF needs to know what's working

- Dynamic, attractive programs that get kids engaged in STEM activities
  - Mobilize communities and partners
- Information that can be used by designers and staff
- Multi-methods and forms of evidence
- Stay open to the unanticipated!

NSF *has* considered all these issues

Examples:

- ACC's Informal Ed and Outreach (2007)
- Friedman Evaluation Framework (2008)
- NRC Learning Science in Inf. Env. (2009)
- Frechtling User-Friendly Handbook (2010)

Need to follow the lead and act!

Indeed, there is a flexible structure

- NSF encourages diversity of
  - Outcomes (“other”)
  - Research and evaluation methods

Show audiences and stakeholders how

- **Projects are true to their objectives**
- **Rigorously across evaluation methods**

Thank You

Menlo Park Headquarters

SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025-3493
650.859.2000

Washington, D.C.

SRI International
1100 Wilson Blvd., Suite 2800
Arlington, VA 22209-3915
703.524.2053

Additional U.S. and international locations

www.sri.com

Components of an ITEST STEM Workforce Development Model

- Strand A: Local context and cultural relevance
- Strand B: STEM skills and knowledge
- Strand C: Partnerships
- Strand D: Career pathways
- Strand E: Youth development
- Strand F: Teacher professional development

technology use and engagement of populations traditionally underrepresented in STEM will be woven into the overall model development work

Process: Three Sessions

- Session 1 (Wednesday @ 1:00–2:00 PM)
  - Share contributions
  - Begin discussion using guiding questions
- Session 2 (Wednesday @ 2:30-4:00 PM)
  - Continue discussion using guiding questions
- Session 3 (Thursday @ 10:00 AM-11:30)
  - Wrap-up and synthesize discussion
  - Prepare for report out
  - Number off for jigsaw conversations

Today's Schedule

- 11:00 Keynote: Dale Dougherty
- 11:45 Break/get lunch in foyer
- 12:00 Panel: What is Evidence?
- 1:00 Breakout Strand Sessions
- 2:00 Break
- 2:30 Continue Strand Sessions
- 4:00 Meet with Program Officers
- 5:00 Poster Setup
- 6:00 Working Dinner/Poster Reception

twitter: #itest Virtual event: http://learningtimesevents.org/itest/

EDC Learning transforms lives.