

# Assessing Readiness for Rigorous Evaluation

---

Cay Bradley & Lauren Scher, Mathematica

February 2, 2021



**OASH**

Office of  
Population Affairs

# Agenda

---

- Introduction
- The “Evaluate” phase – what evaluation activities make the most sense for your interventions?
- Planning a rigorous evaluation

# Today's Speakers

---



**Cay  
Bradley**  
Principal  
Researcher



**Lauren  
Scher**  
Senior  
Researcher

## Hiking ... a metaphor for developing evidence

---



## Flat Lands ... Exploring and Developing

---

- Develop (and disseminate) an understanding of
  - The problem (including need and demand)
  - The target population
  - Available services and gaps
  - Available interventions
  - Support from community
- Develop (and document) innovations/interventions
  - Define innovation/intervention
  - Design supports for implementation – including logic model, fidelity and outcome measures, and continuous quality improvement system

## Pros of Spending Time in Flat Lands

---

- Is your program really an innovation?
- Will your target population sign-up? Attend?
- Can you find partners to implement the program?
- Do you need to provide supports to partners?
- Do facilitators implement with fidelity?
- What supports do facilitators need?

## Rolling Hills ... Test and Refine

---

- Demonstrate and document
  - Feasibility – focus on PROCESS not outcomes
    - ✓ Recruitment
    - ✓ Data collection procedures
    - ✓ Integration into existing systems
    - ✓ Adaptability
    - ✓ Implementation
  - Outcomes
    - ✓ Measurement
    - ✓ Possibly document change
- Repeat more than once



## Pros of Mastering Rolling Hills

---

- Have the “right” measures/instruments
  - Know that program can change key outcomes
- Develop CQI processes to ensure delivering high-quality
- Ensure all content and activities resonate with participants and facilitators
- Be confident target population will enroll and participate
- Gather evidence to secure necessary funding to support expanding program and/or expanding evaluation services

## Snow-Covered Mountains – Evaluate Phase

---

- Focus on building comprehensive evidence
- Different designs
  - Single group pre/post, QED, or RCT
  - Pilot or larger scale
- Decisions driven by
  - Readiness (documentation/dissemination of evidence)
  - Resources (money, time, commitment, capability)

## Rushing the “hike”

---

- Big investment and limited learning
- Testing something other than intended innovation
- Lack of evidence, or evidence it doesn’t work
- Frustration

## Starting the climb: moving to “Evaluate” phase

---



## Establishing merit: moving to larger pilot

---

- Pilot testing in the Evaluate phase: less controlled environment – in more or different settings with more participants
- Research questions a larger pilot study can answer
  - Is there community need and demand for the program and is the innovation aligned with those needs?
  - Is there clear support and satisfaction from the community and participants/target population?
  - Is the innovation a good fit?
  - Can the innovation be implemented as intended (e.g., with fidelity, high quality)?
  - Do outcomes in the logic model improve? (proximal outcomes)

## Methods and data during the pilot test phase

---

- Use mixed methods (qualitative and quantitative data)
  - Pre/post outcomes from surveys or admin records
  - Key informant interviews, focus groups with a range of interested parties
  - Document reviews
  - Program observations
  - Fidelity and other implementation data (e.g., dosage, participation, reach, participant engagement, quality measures)
- Developing data collection and tracking tools, navigating IRB, and data analysis will help set the stage for a randomized controlled trial (RCT) or quasi-experimental design study (QED) or another method for assessing impacts

## Documenting the pilot experience

---

- Regardless of the outcome of the pilot, you should be prepared to document
  - Qualitative and quantitative findings answering your key research questions
  - Contextual factors that may support or act as barriers to implementation
  - Lessons learned and suggestions for program improvement
  - Recommendations for next steps
    - ✓ Back to “Test and Refine”?,
    - ✓ Additional pilot testing with a broader or different population?
    - ✓ Readiness for impact evaluation

## When is there merit for an impact evaluation?

- ✓ Environmental scan suggests need and shows community lacks similar programming.
- ✓ Participants (and/or caregivers) reported being satisfied with program.
- ✓ Interested parties found it compelling and a good fit.
- ✓ The innovation is clearly defined and implemented with fidelity and quality.
- ✓ Near-term outcomes improved (and logic model suggests longer term behavioral outcomes will too).



## Think about future impact evaluation design

---

- Impact designs can include RCTs, QEDs, interrupted time series, regression discontinuity designs, stepped wedge, single case
- Today we will talk about planning for RCTs and QEDs
  - Require: control or comparison group (either no treatment or an alternative treatment)
  - Require: large enough sample size to detect statistically significant effects
- We assume that most of you will not consider an impact evaluation during this grant period (jumping in too quickly likely would lead to small, non-significant impacts)

## Developing a summative (impact) evaluation plan

---

- ✓ Clearly specify research questions
- ✓ Choose a study design to answer research questions
- ✓ Prepare to demonstrate baseline equivalence
- ✓ Clarify the contrast you plan to test
- ✓ Determine an appropriate sample size
- ✓ Design study to avoid confounding factors
- ✓ Plan for recruitment of eligible sample
- ✓ Consider strategies for obtaining consent
- ✓ Specify data collection and analysis procedures
- ✓ Plan for disseminate to a range of stakeholders

## Clearly specify research questions

---

- What do you want to learn?
  - Overall program impact?
  - Testing a component or comparing alternative components?
- Study design driven by your **research questions**
  - Align with logic model
  - Specify primary and secondary outcomes of interest, time frame, and subgroups of interest
- *EXAMPLE RQ: “What is the impact of ‘INNOV’ on sexual initiation rates six months following the program?”*

## Randomized controlled trials (RCTs)

---

- RCTs use a random mechanism to assign sample to two or more groups
  - Randomization can occur at the individual-level (youth) or at a group or cluster level (e.g., schools, teachers or classrooms)
  - Because of randomization, groups should be similar on both observed and unobserved characteristics
  - In a well-implemented RCT, we can feel more confident that differences between groups on outcomes are due to the intervention

## Assessing the feasibility of an RCT design

---

- Are you confident the program will make a difference on outcomes that your partners care about?
- Is random assignment feasible?
- Are partners on board with random assignment?
  - Requires collaboration, creativity, and clear communication of the value of the study design and expectations for maintaining the design

## Quasi-experimental designs (QEDs)

---

- QEDs use a less rigorous, non-random mechanism to form the treatment and comparison groups
  - Even if groups appear similar on observed characteristics such as age or race, they may be different in some unobserved way
  - Differences in outcomes between groups may be unrelated to the intervention.
  - Are good options when RCTs are not feasible, but care should be taken to ensure that groups similar.

# Design considerations in an impact study

---

- Prepare to demonstrate baseline equivalence
  - Specified characteristics (e.g., age, race, ethnicity, gender)
  - Baseline measures of the outcome of interest
- Clarify the contrast you plan to test
  - Make sure the contrast will answer your research question
  - Plan to measure the contrast
- Determine an appropriate sample size
  - You can use assumptions from your pilot / prior research
- Design study to avoid confounding factors
  - Need more than 1 treatment / 1 comparison cluster (e.g., school)
  - Plan for same data collection procedures (e.g., delivery, timing)

## Plans for implementing a rigorous study

---

- Plan for recruitment of eligible sample
  - Determine a realistic timeline for recruitment
  - Consider study eligibility requirements
- Determine strategies for obtaining consent / assent
- Specify data collection procedures
  - What data do you need? Who will collect it?
  - Develop or refine data collection instruments / protocols
  - Develop plans for following up with sample (contact info, outreach, incentives, IRB requirements)
- Develop plan for companion process and implementation evaluation – **strongly recommended!!!**

## Pre-plan for analysis and dissemination

---

- Specify planned analytic methods aligned to your model
- Plan to disseminate to a range of stakeholders throughout the course of your study



# Resources

---

- [Impact evaluation toolkit](#)
- YOUR EVAL TA LIAISON!
  - Review and provide feedback on evaluation plans
  - Help consider evaluation design options
  - Provided suggestions on measures to use



**OASH**

Office of  
Population Affairs

## Q & A

---

## Upcoming webinars

---

- Systems and network analyses      March 24 @ 2pm ET
- Core components                         June 8 @ 2pm ET

# Thank you!!!

---

- Cay Bradley: [cbradley@mathematica-mpr.com](mailto:cbradley@mathematica-mpr.com)
- Lauren Scher: [lscher@mathematica-mpr.com](mailto:lscher@mathematica-mpr.com)