



KOICA Executive Education Course in
Evaluating Social Programs
22-26 February 2016
Welcome

Lina Marliani
Executive Director
J-PAL Southeast Asia





Course Introduction



Course Leaders

- Thomas Chupein
- John Floretta
- Lina Marliani
- Rohit Naimpally
- Ariella Park
- Hira Siddiqui

Course Overview

1. Introducing Randomized Impact Evaluations (Thomas Chupein)
2. Outcomes, Impact, and Indicators (Rohit Naimpally)
3. Why Randomize? (Thomas Chupein)
4. How to Randomize (Rohit Naimpally)
5. Sampling and Sample Size (Rohit Naimpally)
6. Threats and Analysis (Lina Marliani)
7. Evaluation from Start to Finish (Lina Marliani)
8. Evidence from Community-Driven Development, Health, and Education Programs (Thomas Chupein and John Floretta)
9. Using Evidence from Randomized Evaluations for Decision-Making and Policy Change (John Floretta)
10. Discussion: Where do RCTs fit in a Good M&E Strategy? (All)

Course Objectives

- Why and when is a rigorous evaluation of social impact needed?
- The common pitfalls of evaluations, and how randomization can help.
- The key components of a good randomized evaluation design
- Alternative techniques for incorporating randomization into project design.
- How do you determine the appropriate sample size, measure outcomes, and manage data?
- Guarding against threats that may undermine the integrity of the results.
- Techniques for the analysis and interpretation of results.
- How to maximise policy impact and test external validity.

Course Agenda

	Monday February 22, 2016	Tuesday February 23, 2016	Wednesday February 24, 2016	Thursday February 25, 2016	Friday February 26, 2016
8:00 – 9:00	Registration and Pre-Course Assessment	Breakfast	Breakfast	Breakfast	Breakfast
9:00 – 10:30	Welcoming remarks Lecture 1: Introducing Randomized Impact Evaluations Thomas Chupein	Lecture 3: Why Randomize? Thomas Chupein	Lecture 5: Sampling and Sample Size Rohit Naimpally	Lecture 7: Evaluation from Start to Finish: Raskin in Indonesia Lina Marliani	Lecture 9: Using Evidence from Randomized Evaluations for Decision-Making and Policy Change John Floretta
10:30 – 10:45	Coffee Break	Coffee Break	Coffee Break	Coffee Break	Coffee Break
10:45 – 12:00	Group work on case study 1: Theory of Change: Women as Policymakers in India Decision on group project	Group Exercise A: Random Sampling Group work on presentation: Indicators	Group work on case study 4: Threats and Analysis: Deworming in Kenya	Group work on presentation: Power and sample size	Discussion: Where do RCTs fit in a good M&E Strategy
12:00 – 1:00	Lunch	Lunch	Lunch	Lunch	Lunch
1:00 – 2:30	Lecture 2: Outcomes, Impact, and Indicators Rohit Naimpally	Lecture 4: How to Randomize Rohit Naimpally	Lecture 6: Threats and Analysis Lina Marliani	Lecture 8: Evidence from Community-Driven Development, Health, and Education Programs Thomas Chupein and John Floretta	Feedback survey
2:30 – 3:00	Coffee Break	Coffee Break	Coffee Break	Coffee Break	
3:00 – 4:00	Group work on presentation: Theory of change, research question	Group Exercise B: Randomization Mechanics	Group work on presentation: Randomization Design	Group work on presentation: Threats and Analysis	Group presentations
4:00 – 5:00	Group work on case study 2: Why Randomize: Learn to Read India	Group work on case study 3: How to Randomize: Combating Corruption in Indonesia	Group Exercise C: Sample Size Estimation	Group work on presentation: Finalize presentation	Closing remarks

Course Structure

- Lectures
- Case Studies
- Exercises

Course Binder

- Course Schedule.....1
- Biographies of J-PAL Lecturers.....5
- Course Material
- Case Study 1: Women as Policymakers.....7
- Case Study 2: Learn to Read Evaluations.....13
- Case Study 3: Combating Corruption.....21
- Case Study 4: Deworming in Kenya.....27
- Exercise A: Random Sampling and Law of Large Numbers.....35
- Exercise B: Mechanics of Randomization37
- Exercise C: Sample Size Estimation.....47
- Group Presentation Guidelines.....63
- Impact Evaluation Glossary.....67

Expectation Survey

J-PAL Executive Education Course
New Delhi, India
July 2015

Participant #:

What are your goals for the course?

In order to gauge how well our course is matching our participants' interests, J-PAL would like to know what participants' goals are going into the course.

Please rank the **4 topics** that most interest you or that you are hoping to learn the most about during the course (indicate your most important goal with a "1", and continue up to "4" in order of decreasing importance):

Understanding what evaluation is and why it is valuable	
Conceptualizing and constructing a logical framework or Theory of Change	
Developing a research question	
Developing indicators to measure outcomes	
Identifying the pros and cons of different types of impact evaluation	
Understanding the basic design of a randomized evaluation	
Randomizing the assignment of a program in the face of practical constraints	
Calculating statistical power/determining sample size	
Selecting an unbiased, representative sample	
Managing an evaluation	
Collecting data	
Using monitoring data to track and improve program implementation	
Understanding and dealing with what can go wrong in a randomized evaluation	
Analyzing data obtained through an evaluation	
Conducting cost-effectiveness analysis	
Making evaluation relevant for policymaking	
Scaling up effective interventions	
Fostering partnerships with researchers for evaluation	

Pre-course Assessment

J-PAL Executive Education Course
New Delhi, India
July 2015

Participant #:

Pre-Course Assessment

Here is a short survey that poses questions about the various topics covered throughout the course. Please answer the questions to the best of your ability. They will provide J-PAL with useful information about how well the course teaches key concepts.

1. Suppose your NGO seeks to launch a chlorine distribution program to improve access to clean water for its beneficiary households. Please indicate which aspect of program evaluation (numbered below) is most appropriate for:

Measuring the effects of chlorine distribution on important health indicators for beneficiary households	
Following whether or not chlorine is actually distributed to beneficiary households	
Constructing a model to describe how chlorine distribution could lead to outcomes of interest (e.g. reduced incidence of diarrhea in children)	
Comparing the health improvements per dollar spent on the chlorine distribution program with health improvements per dollar spent on other clean water programs	
Identifying the prevalence of diarrhea and the subpopulation that does not currently have access to clean water	

1. Needs Assessment
 2. Program Theory Assessment
 3. Process Evaluation
 4. Impact Evaluation
 5. Cost Effectiveness Analysis
-

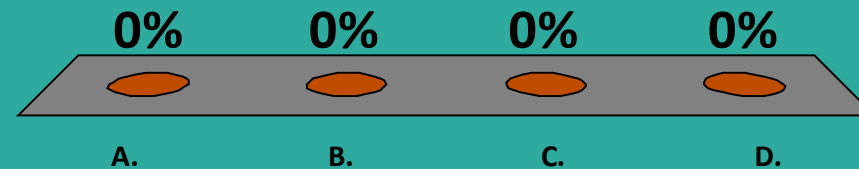
2. Define the counterfactual:

Course “Clickers”

- Everyone gets one
- Each clicker has number on back that matches the number assigned to you on the participant list
- Please hold on to them and turn them in at the end of each day

Course "Clickers": Have you used these before?

- A. Yes
- B. No
- C. Something similar
- D. Something different





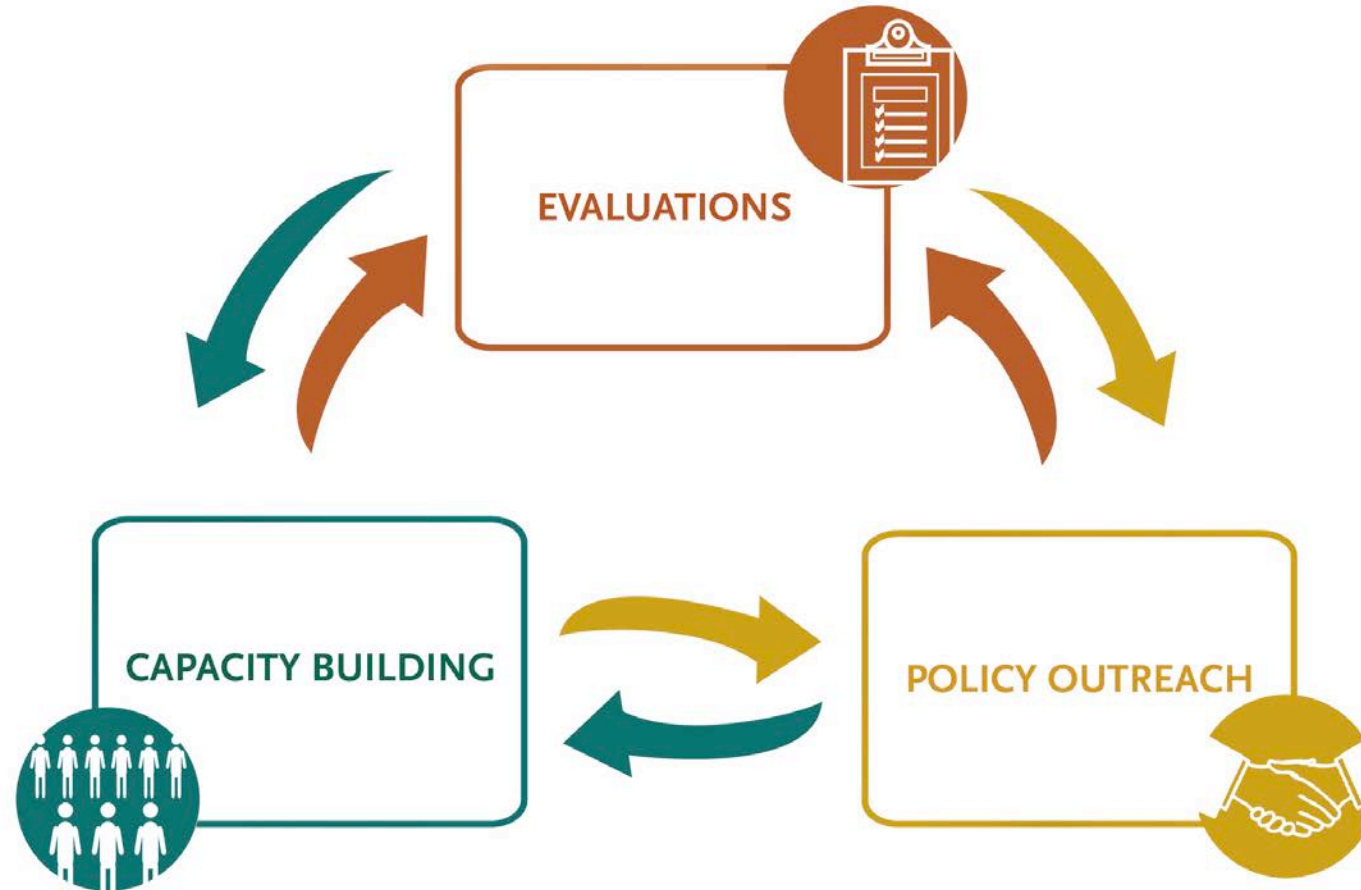
Introduction to J-PAL



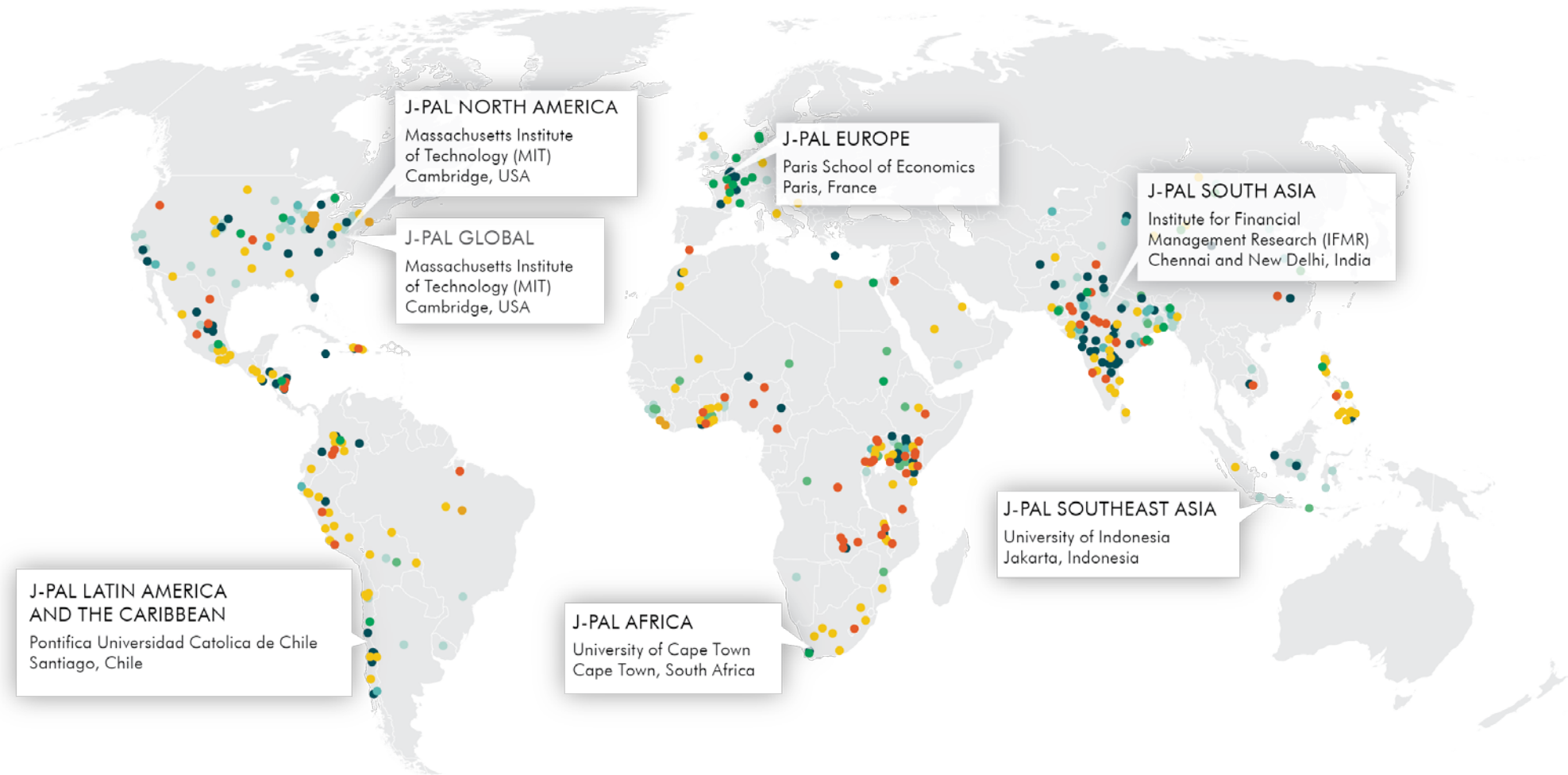
J-PAL's mission is to ensure that policy is informed by evidence and research is translated into action



J-PAL's network of 131 professors use randomized evaluations to inform policy



J-PAL has 7 offices and 700 ongoing and completed projects in 64 countries



J-PAL's activities are organized into eight sectors



Agriculture



Crime



Education



Environment
and Energy



Finance and
Microfinance



Health



Labor Markets



Political Economy
and Governance

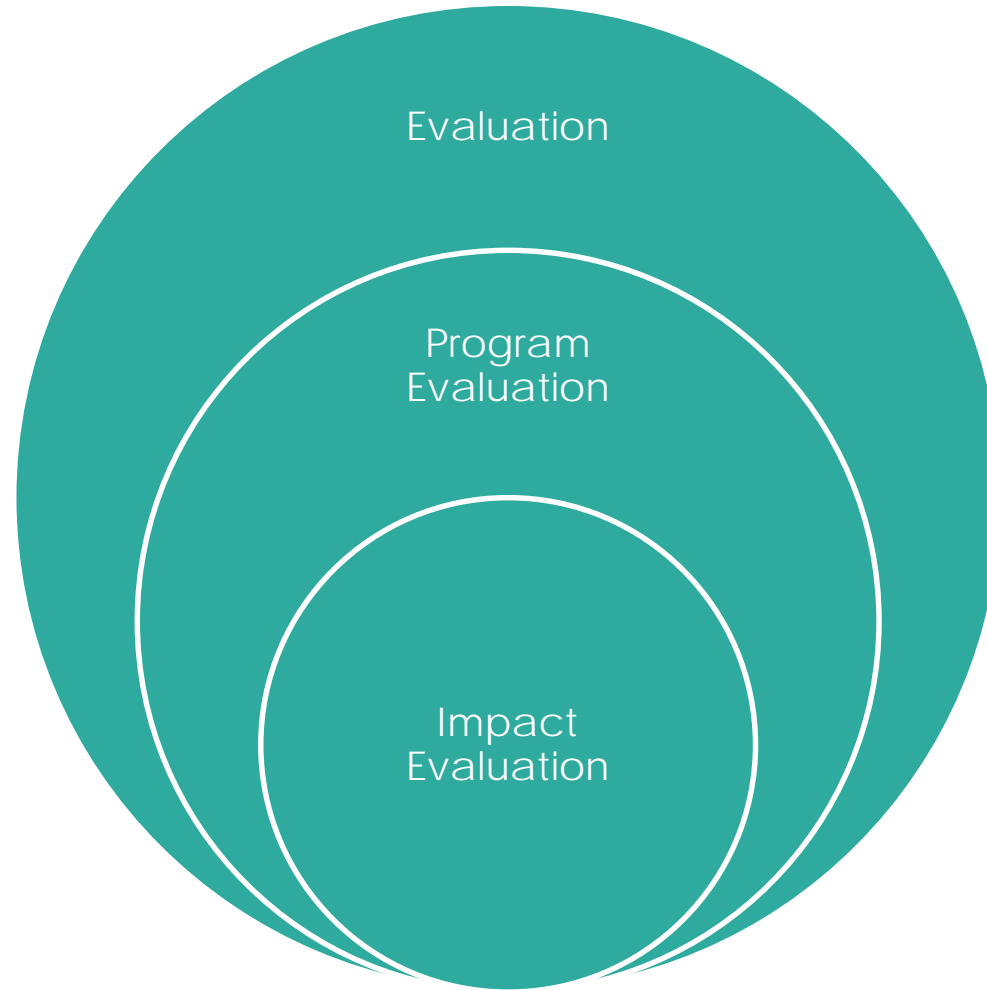


Introducing Randomized Impact Evaluations

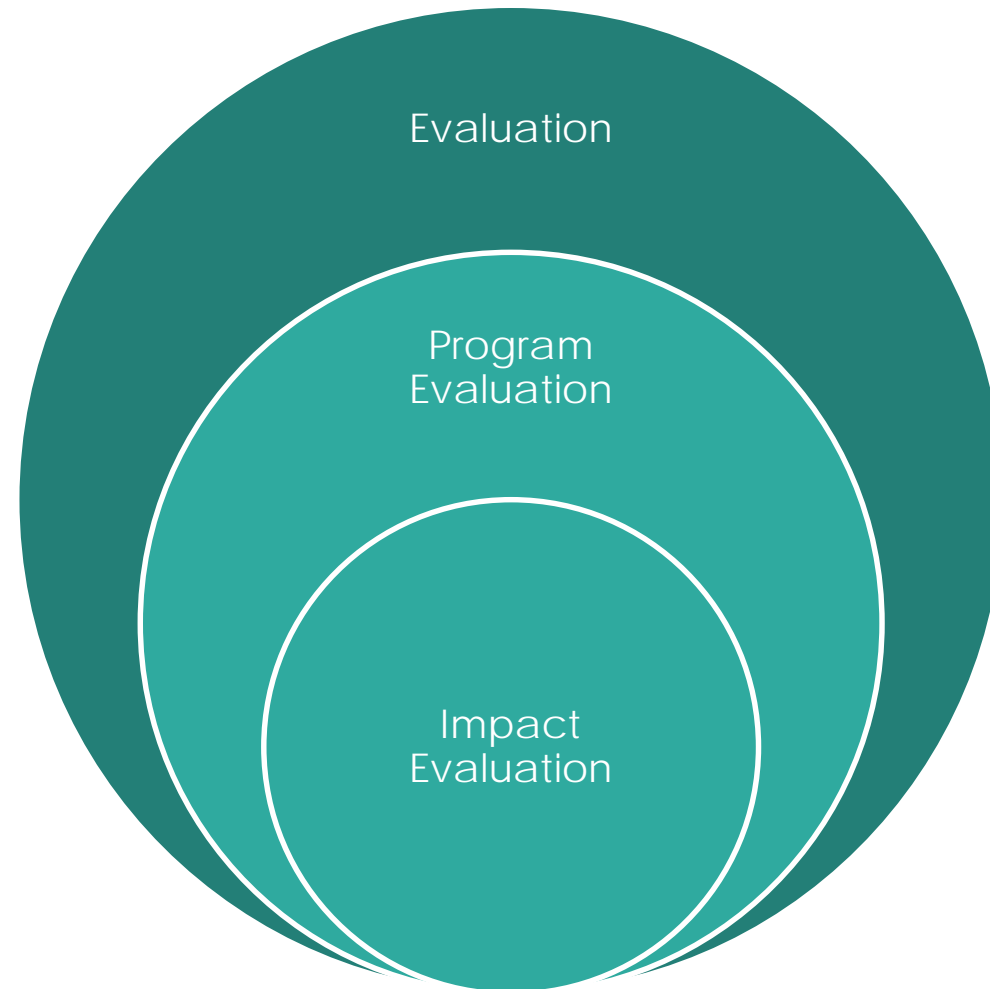
Thomas Chupein
Policy Manager
J-PAL Global at MIT



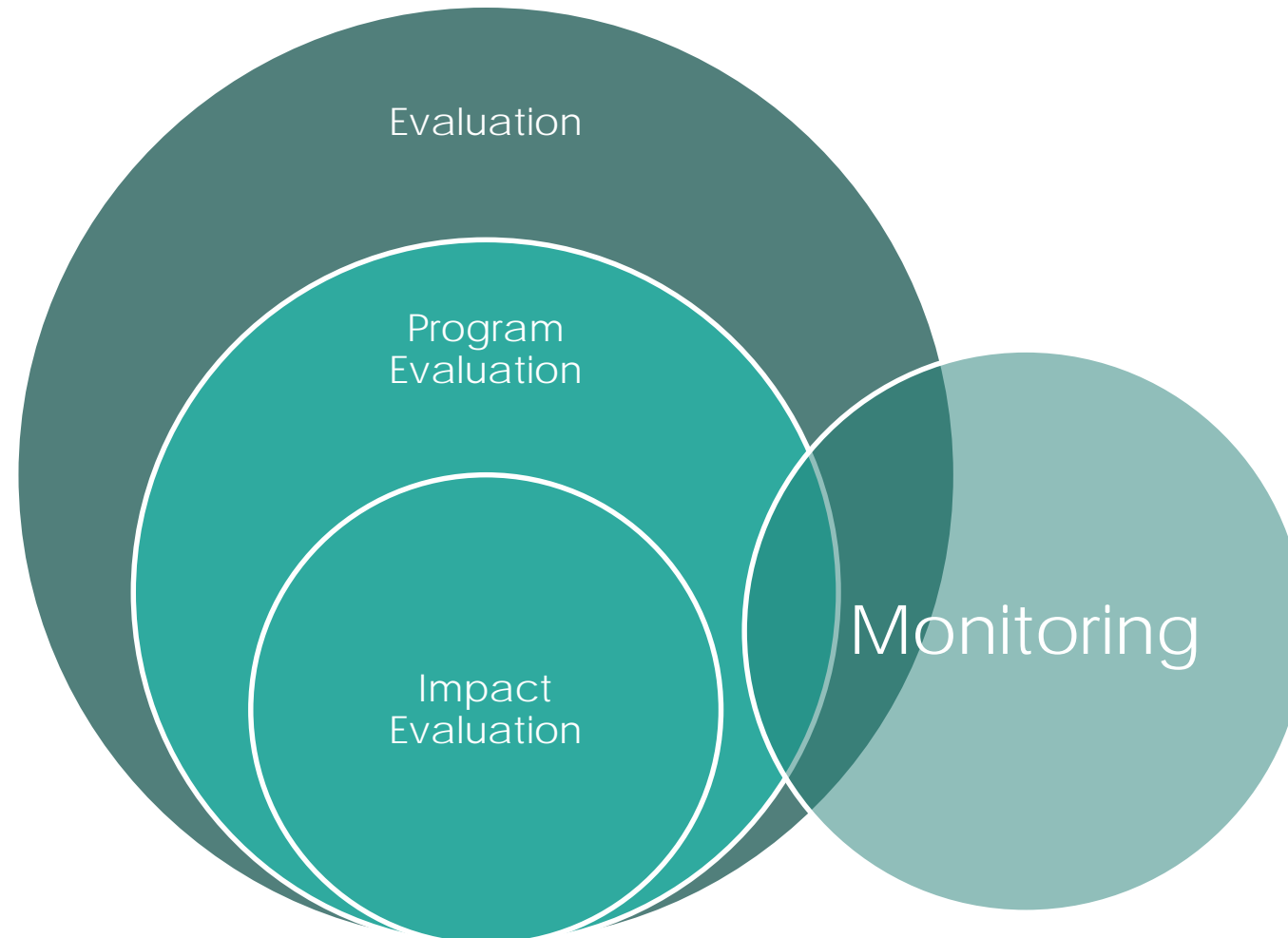
What is Evaluation?



Program Evaluation

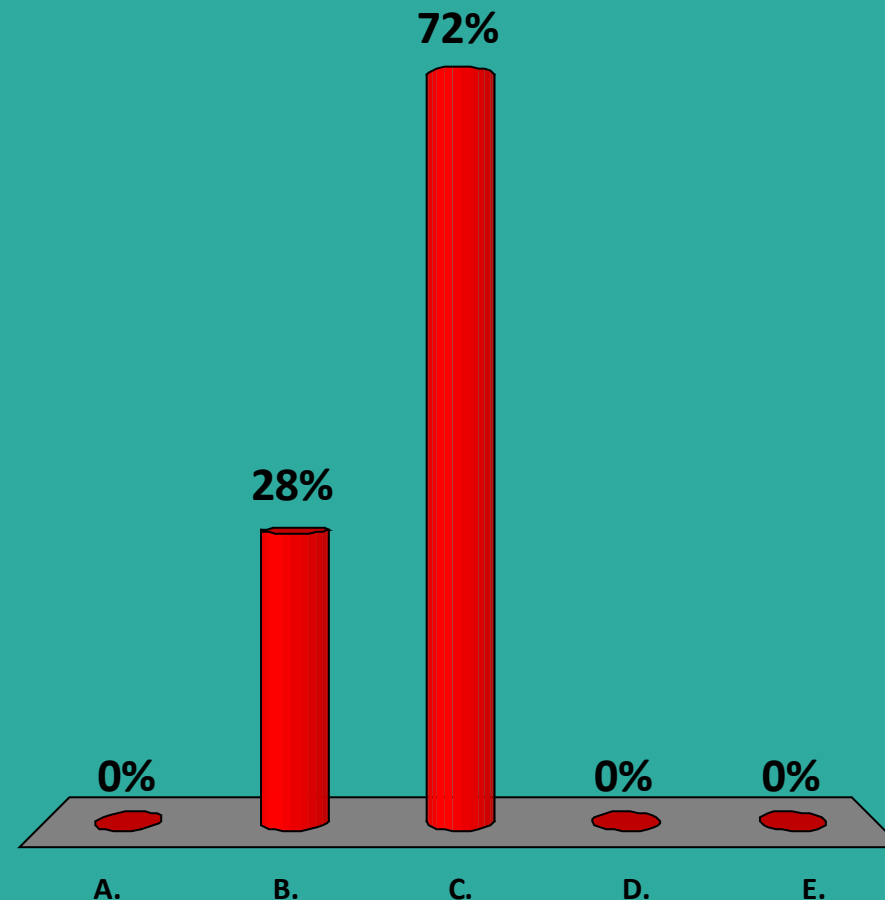


Monitoring and Evaluation

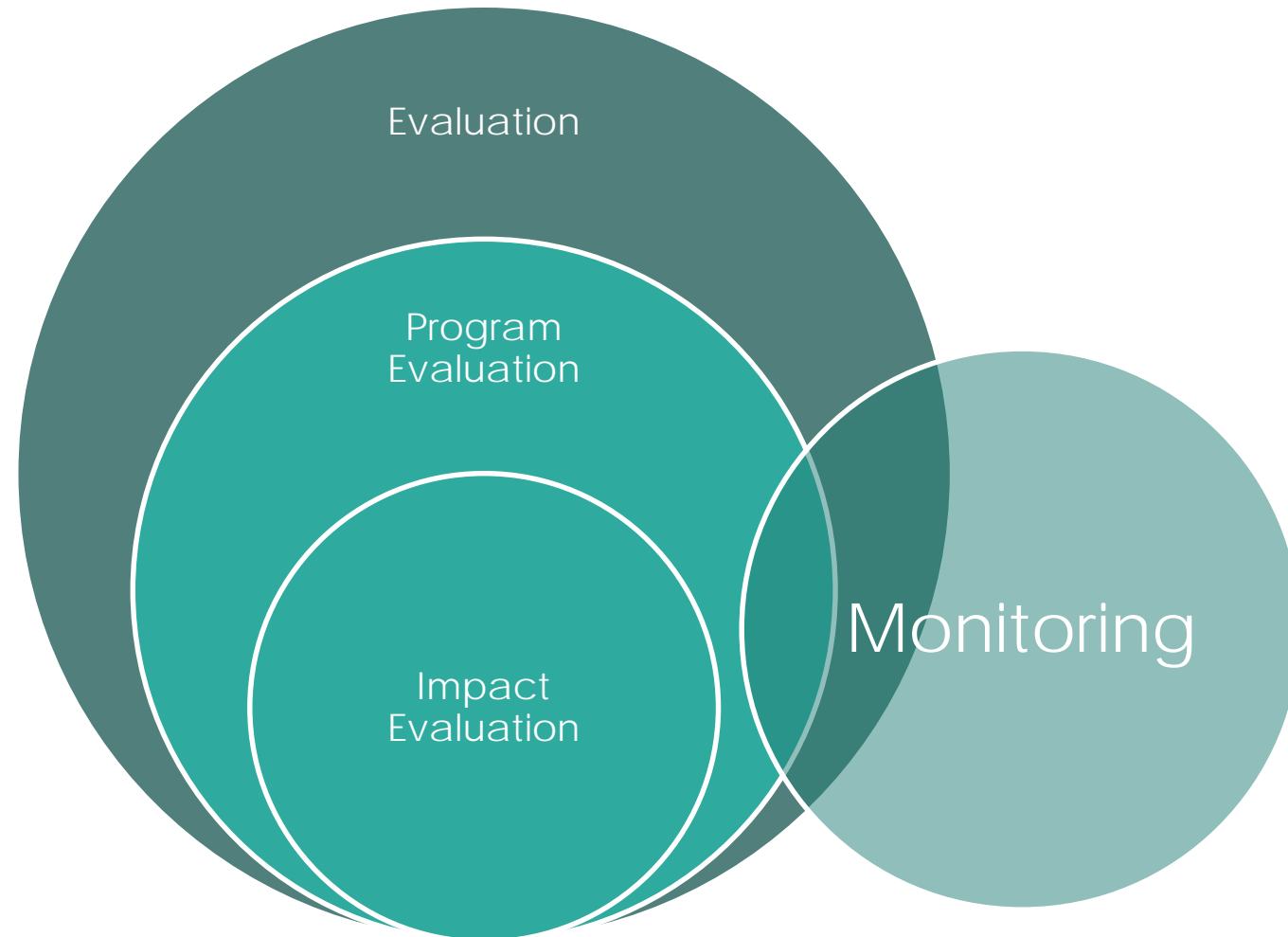


What's the difference between: Monitoring and Evaluation

- A. Nothing. They are different words to describe the same activity
- B. Monitoring is conducted internally, Evaluation is conducted externally
- C. Monitoring is for management, Evaluation is for accountability
- D. Don't know
- E. Other



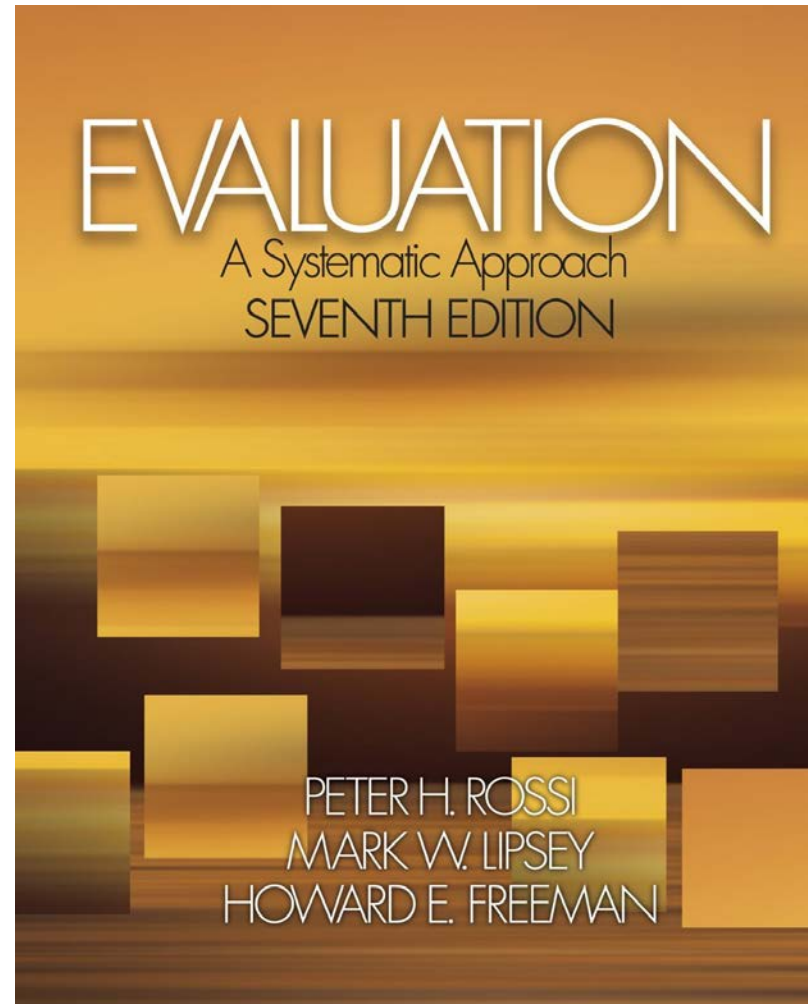
Program Evaluation



5 Components of Program Evaluation

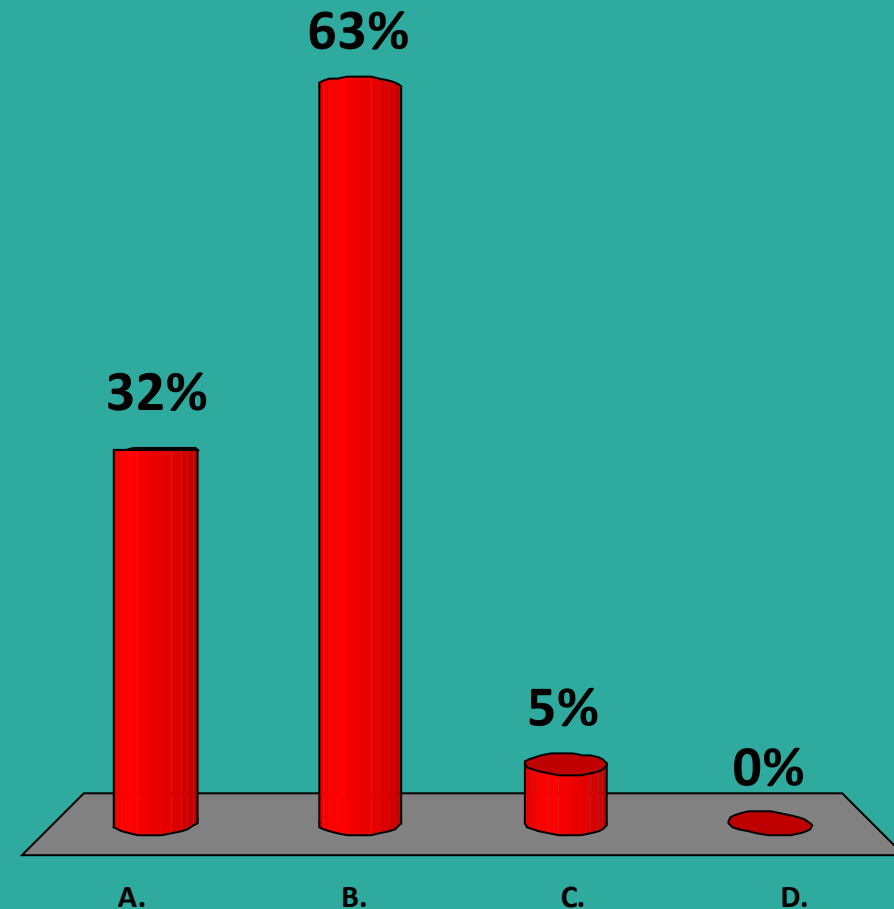
- Needs Assessment
 - Program Theory Assessment
 - Process Evaluation
 - Impact Evaluation
 - Cost Effectiveness
- What is the problem?
 - How, in theory, does the program fix the problem?
 - Does the program work as planned?
 - Were its goals achieved?
The magnitude?
 - Given magnitude and cost, how does it compare to alternatives?

Rossi, Lipsey and Freeman



Evaluation should usually be conducted:

- A. Externally and independent from the implementers of the program being evaluated
- B. Externally and closely integrated with program implementers
- C. Internally
- D. Don't know



How can Impact Evaluation Help Us?

- Surprisingly little rigorous evidence on what works
- Can do more with given budget with better evidence
- If people knew donor funds are used for effective programs, this could help increase resources for anti-poverty programs
- Instead of asking “do aid/development programs work?” should be asking:
 - Which work best, why and when?
 - How can we scale up what works?

Programs and their Evaluations: Where do we start?

Intervention

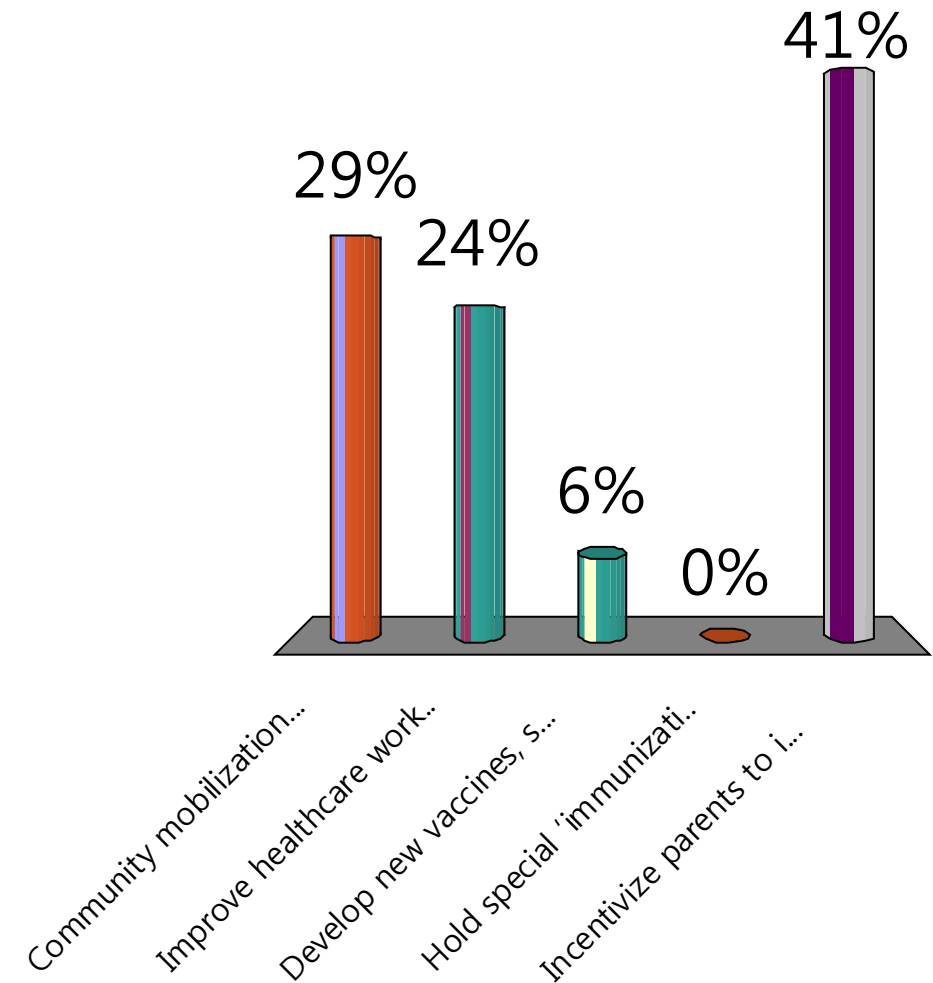
- Start with a problem
- Verify that the problem actually exists
- Generate a theory of why the problem exists
- Design the program
- Think about whether the solution is cost effective

Program Evaluation

- Start with a question
- Verify the question hasn't been answered
- State a hypothesis
- Design the evaluation
- Determine whether the value of the answer is worth the cost of the evaluation

What do you think is the most cost-effective way to increase immunization rates?

- A. Community mobilization campaign
- B. Improve healthcare worker attendance
- C. Develop new vaccines, such as pneumococcal
- D. Hold special 'immunization camps'
- E. Incentivize parents to immunize their children



An Example

Child Immunization



Identifying the problem

1. Needs Assessment



The Need

- Every year, between 2 and 3 million people die from vaccine-preventable diseases
- In India, only 54% of 1-2 year olds receive the basic package of immunizations
- In rural Rajasthan, this rate falls to 22%



The Problem

- In India, immunizations are offered for free, but the immunization rate remains low
- Average household is within 2 kilometers of the nearest clinic
- High absenteeism at government health facilities – 45% of Auxiliary Nurse Midwives are absent on any given workday

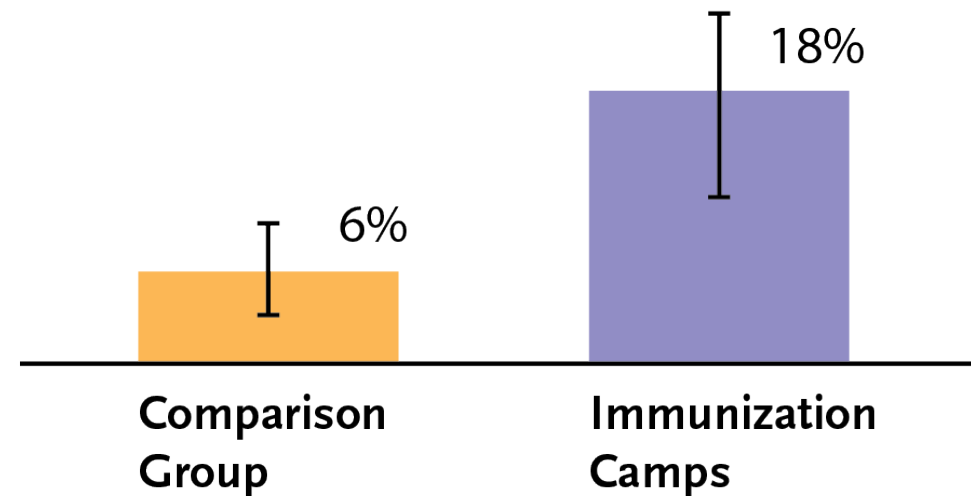
But is Supply the Entirety of the Problem?

- Is there cultural resistance or distrust in public health institutions?
- Parents may not be able to afford to take a day off of work
- People may not value immunizations: short-term cost for long-term (and invisible) benefits

One Potential Solution: Solve the Supply Problem

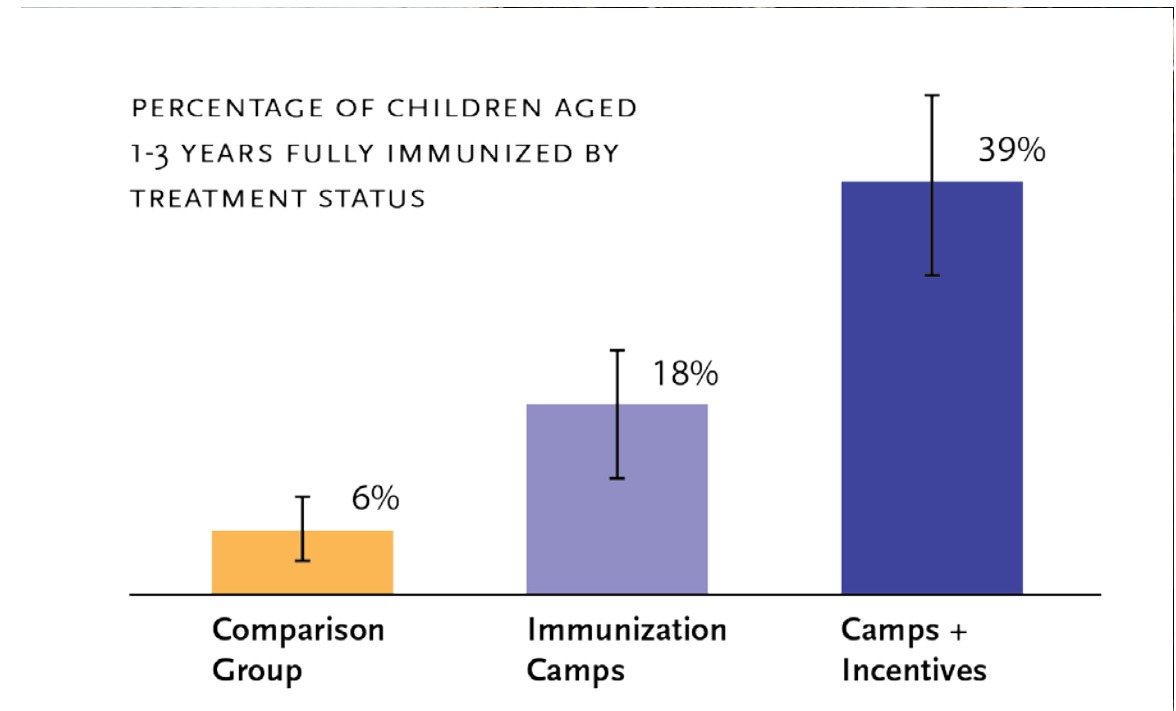
- Reliable NGO-provided monthly immunization camps at the village level tripled rates of full immunization.

PERCENTAGE OF CHILDREN AGED 1-3 YEARS FULLY IMMUNIZED BY TREATMENT STATUS



Another Potential Solution: Solve the Demand Problem

- Alongside reliable NGO-provided monthly immunization camps, offer mothers a small incentive to bring their child
- When coupled with in-kind incentives (raw lentils and metal plates for completing immunization), full immunization rates increased six-fold



Behavioral theory on use of incentives



Blueprint for Change

2. Program Theory Assessment



Program Theory Assessment

- Theory of Change
 - Logical Framework (LogFrame)
 - Results Framework
 - Outcome Mapping
- Causal chain
 - Causal model
 - Cause map
 - Impact pathways
 - Intervention theory
 - Intervention framework
 - Intervention logic
 - Investment logic
 - Logic model
 - Outcomes chain
 - Outcomes hierarchy
 - Outcome line
 - Program logic
 - Program theory
 - Programme theory
 - Results chain
 - Theory-based evaluation
 - Theory-driven evaluation
 - Theory-of-action

What is a Theory of Change?

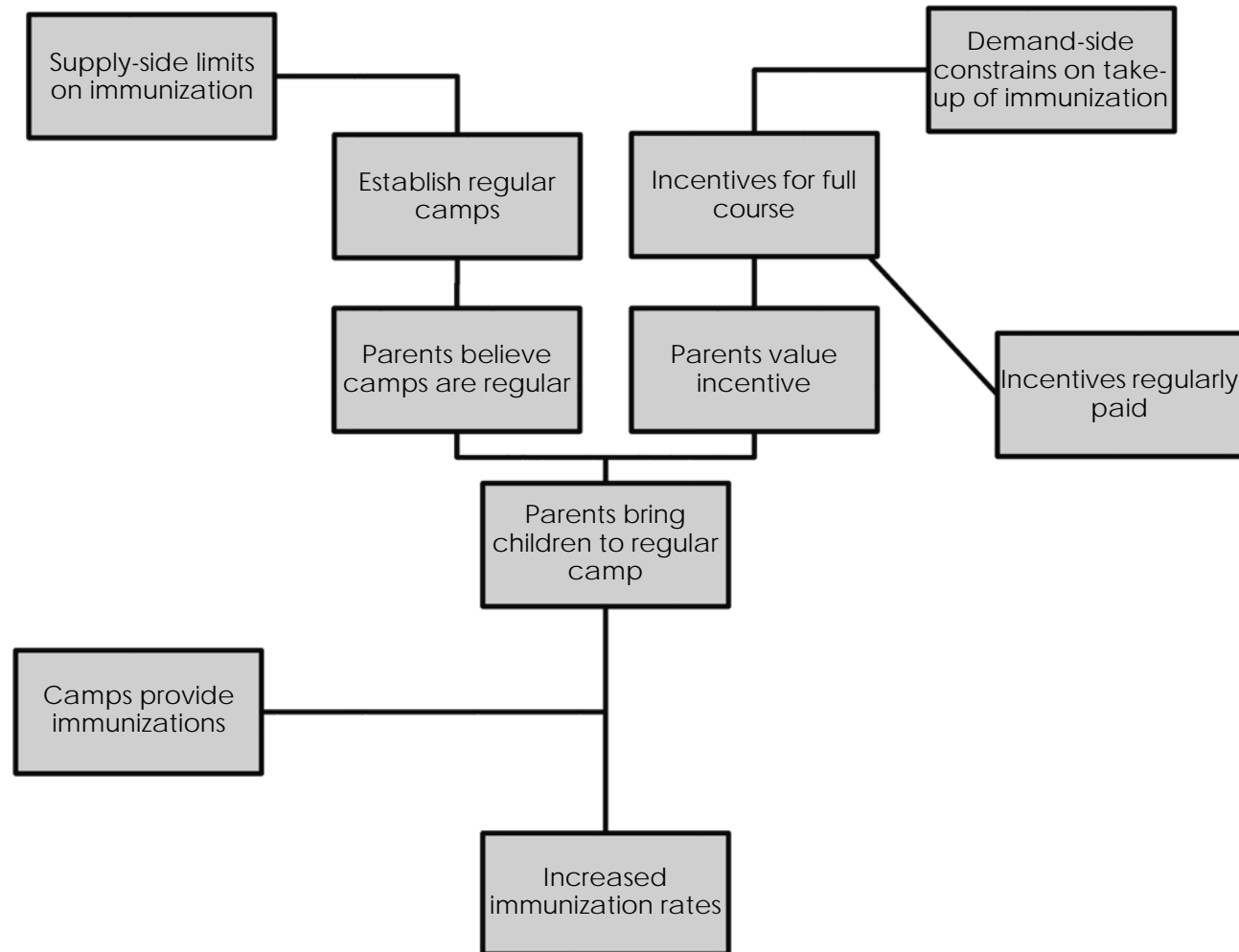
“A theory of change is a road map of where we are going (results) and how we are getting there (process)”

Causal Hypothesis

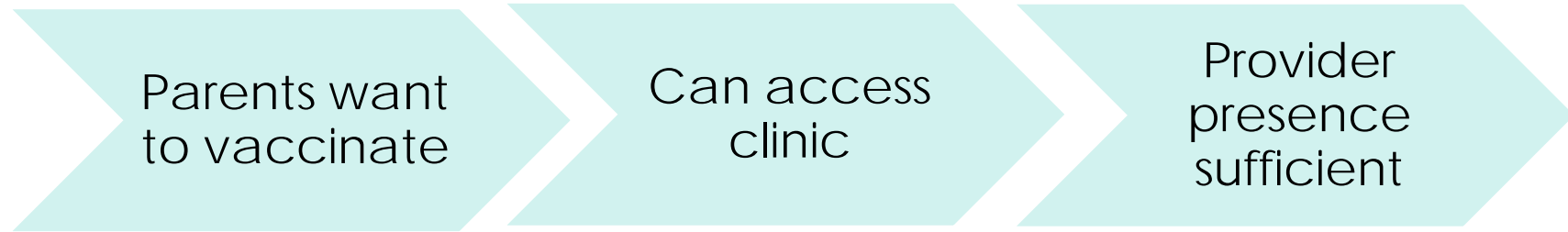
Q: How do I expect results to be achieved?

A: *If [inputs] and [activities] produce [outputs] this should lead to [outcomes] which will ultimately contribute to [goal]*

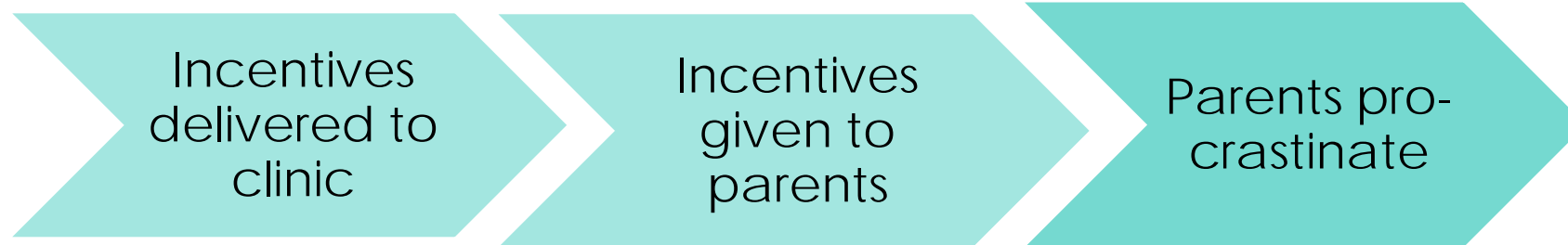
Immunization Program Theory of Change



Immunization Program Theory of Change



Do basic conditions hold locally?



Local logistics critical

Evidence on behavioral bias

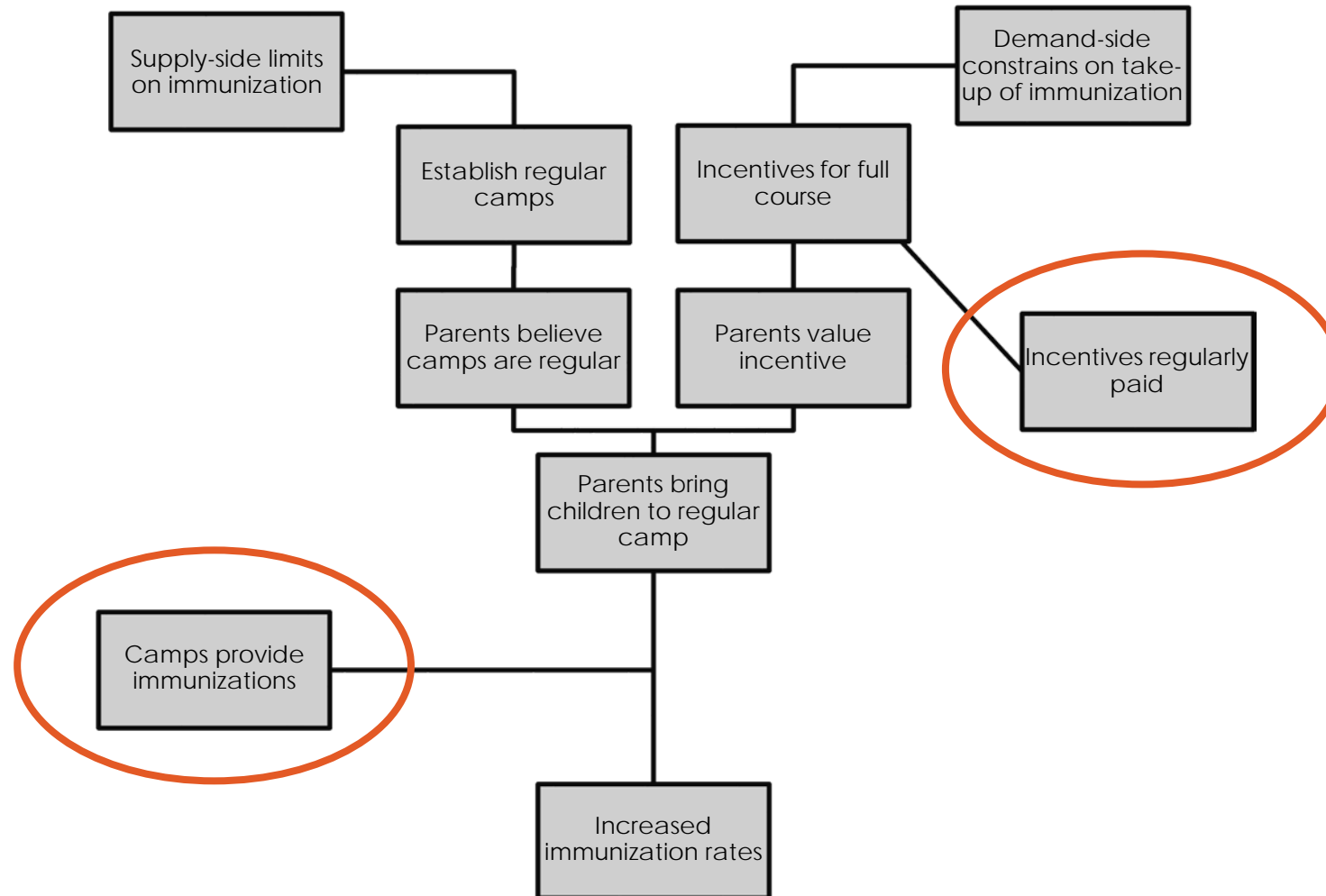


Impact

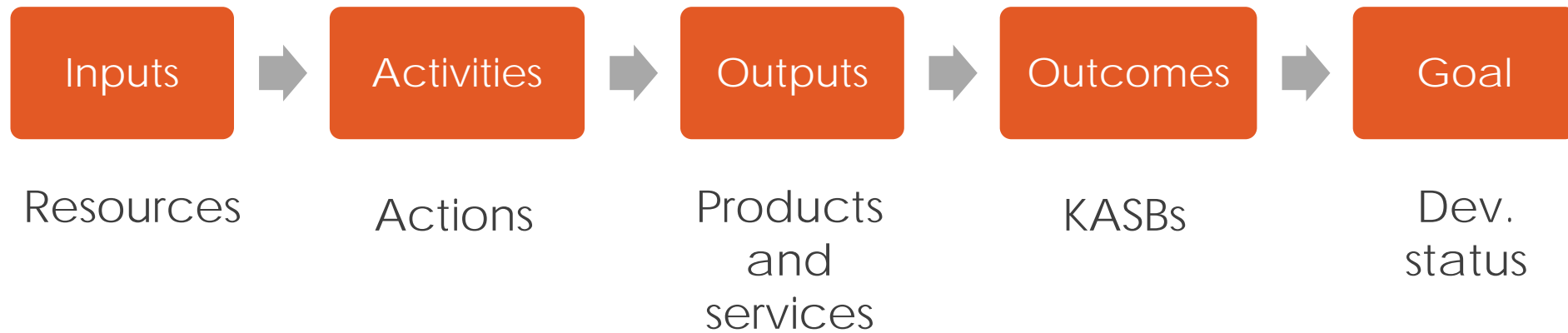
Assumptions

Necessary and positive external conditions that should be in place for the chain of cause and effect (in an intervention) to go forward

Immunization Program Theory of Change



Results Levels



Log Frame

	Objectives Hierarchy	Indicators	Sources of Verification	Assumptions / Threats
Impact (Goal/ Overall objective)	Increased immunization	Immunization rates	Household survey	Adequate vaccine supply, parents do not have second thoughts
Outcome (Project Objective)	Parents attend the immunization camps repeatedly	Follow-up attendance	Household survey; Immunization card	Parents have the time to come
Outputs	Immunization camps are reliably open; Incentives are delivered	Number of kg bags delivered; Camp schedules	Random audits; Camp administrative data	Nurses/assistants will show up to camp and give out incentives properly
Inputs (Activities)	Camps + incentives are established	Camps are built, functional	Random audits of camps	Sufficient materials, funding, manpower

Needs assessment



Impact evaluation

Process evaluation

Source: Roduner, Schlappi (2008) Logical Framework Approach and Outcome Mapping, A constructive Attempt of Synthesis

Theory of Change: Product or Process?

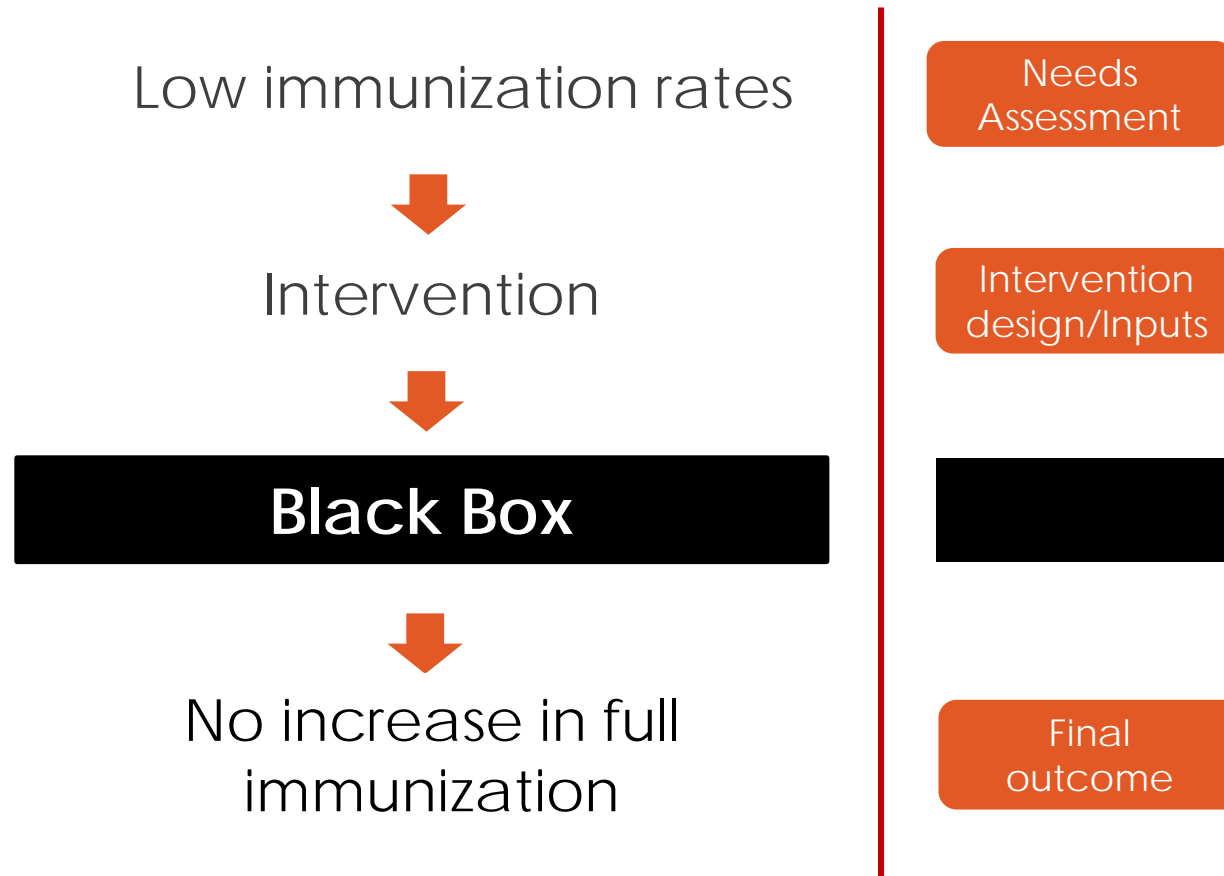
“Theory of change thinking is a habit not a product.”

Making the program work

3. Process Evaluation



Solving the Black Box Problem

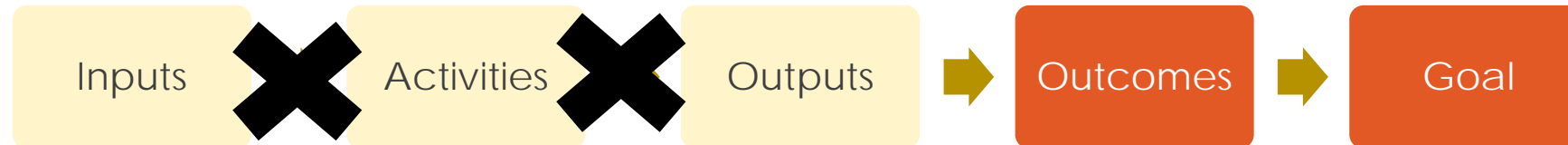


Identifying Implementation Failure vs. Theory Failure

Successful intervention



Implementation Failure



Theoretical Failure



Process Evaluation

- **On the supply side**
 - Logistics
 - Management
- **On the demand side**
 - Assumption of knowledge, preferences
 - Assumptions of response

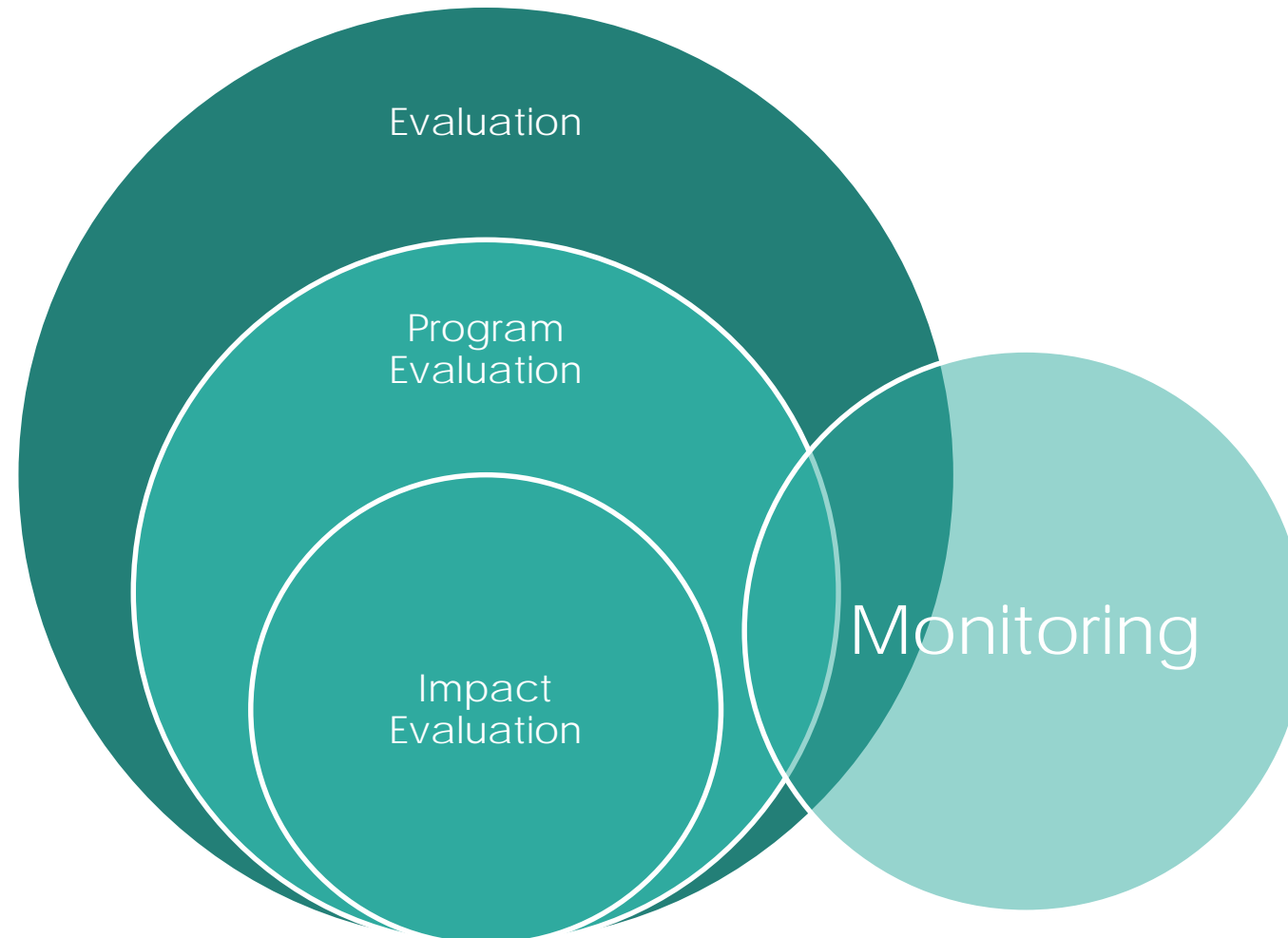
Process Evaluation: Supply-Side

- Establish camp
 - Hiring nurses and administrators
 - Installing temporary camp site
 - Procuring vaccines and other medical supplies
- Organize incentive scheme
 - Identify viable incentive
 - Purchase lentils and dinner plate sets

Process Evaluation: Demand-Side

- Do parents visit the camps?
- Do they come back?

Monitoring and Evaluation



Log Frame

	Objectives Hierarchy	Indicators	Sources of Verification	Assumptions / Threats
Impact (Goal/ Overall objective)	Increased immunization	Immunization rates	Household survey	Adequate vaccine supply, parents do not have second thoughts
Outcome (Project Objective)	Parents attend the immunization camps repeatedly	Follow-up attendance	Household survey; Immunization card	Parents have the time to come
Outputs	Immunization camps are reliably open; Incentives are delivered	Number of kg bags delivered; Camp schedules	Random audits; Camp administrative data	Nurses/assistants will show up to camp and give out incentives properly
Inputs (Activities)	Camps + incentives are established	Camps are built, functional	Random audits of camps	Sufficient materials, funding, manpower

Needs assessment



Impact evaluation

Process evaluation

Source: Roduner, Schlappi (2008) Logical Framework Approach and Outcome Mapping, A constructive Attempt of Synthesis

With process evaluation, we learn...

- Was the program implemented as planned?
- Did people respond as expected?

- If so, then what happened to immunization rates?

Measuring how well the program worked

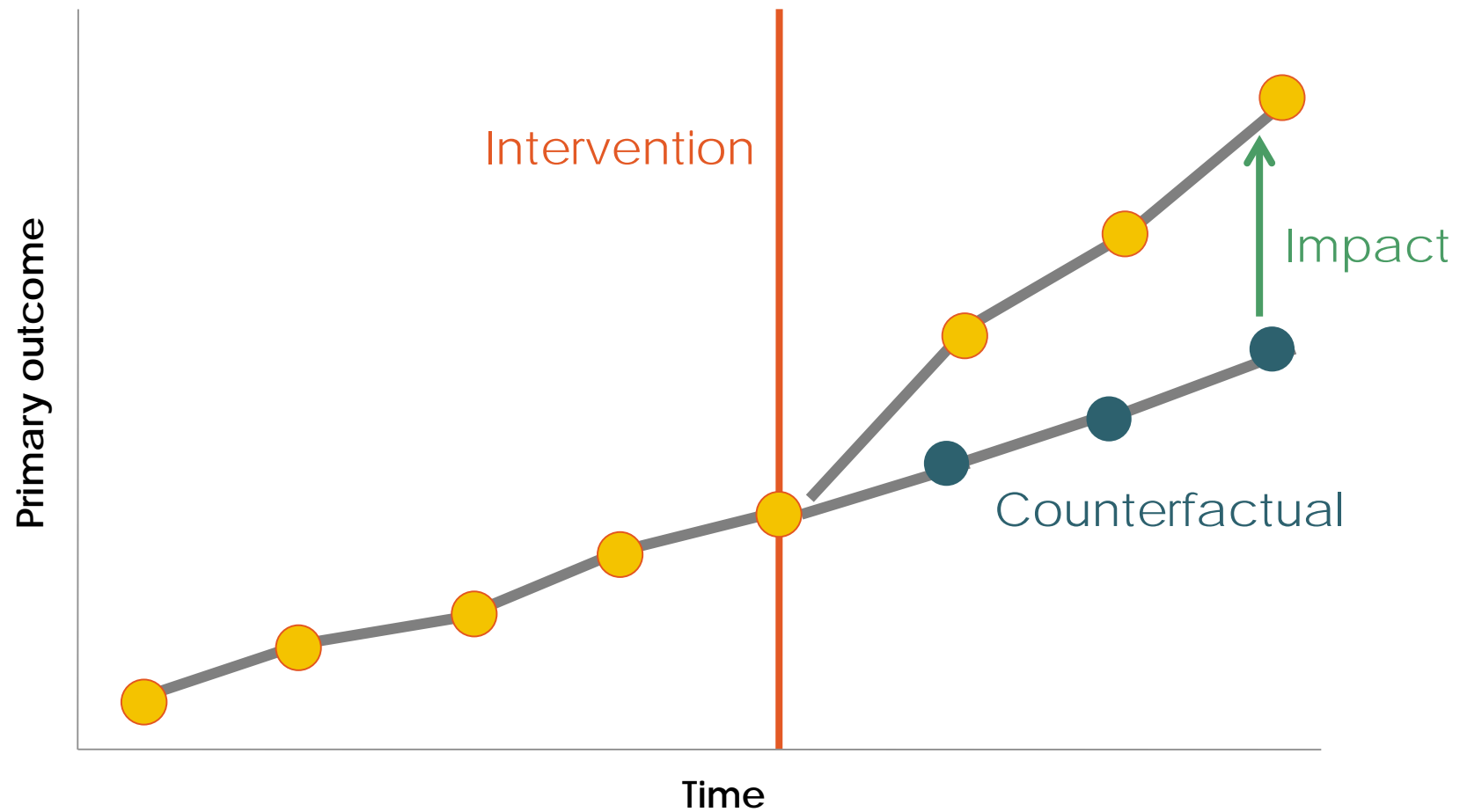
4. Impact Evaluation



Did we Achieve our Goals?

- Primary outcome (impact): did camps (or camps + incentives) raise children's full immunization rates?
- Also distributional questions: for example, what was the impact of incentives for households by level of income?

What is Impact?



How to measure impact?

- We need to know the counterfactual (i.e. What would have happened in the absence of the program?)
- Take the difference between
 - what happened (with the program) and
 - what would have happened (without the program)
- This yields the impact of the program

Constructing the Counterfactual

- Counterfactual is often constructed by selecting a group not affected by the program
- Randomized method:
 - Use random assignment of the program to create a control (comparison) group which mimics the counterfactual.
- Non-randomized method:
 - Argue that a certain excluded group mimics the counterfactual.

How impact differs from process?

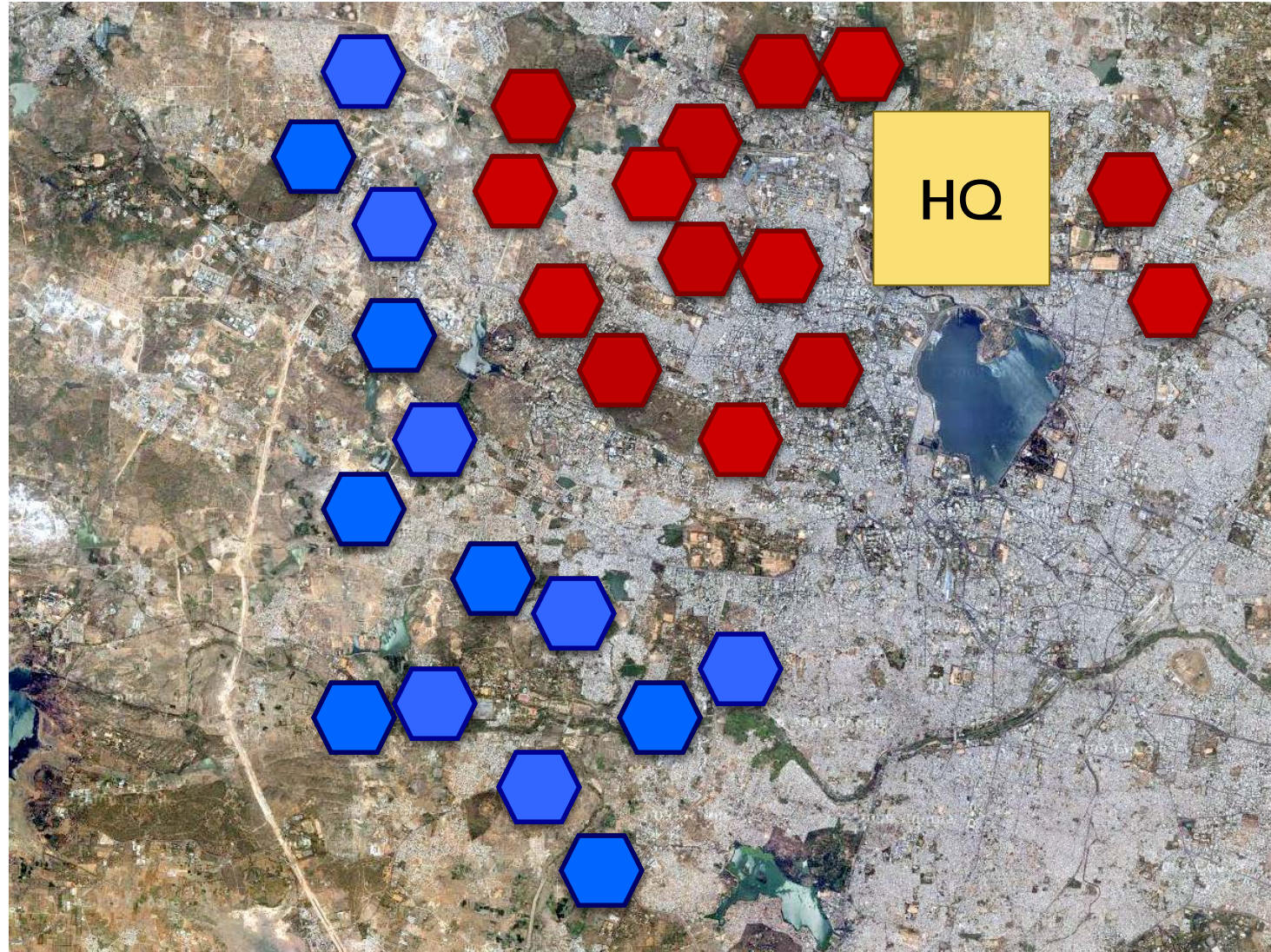
- When we answer a *process* question, we need to describe **what happened**.
- When we answer an *impact* question, we need to compare **what happened** to **what would have happened without the program**

Randomized Evaluation



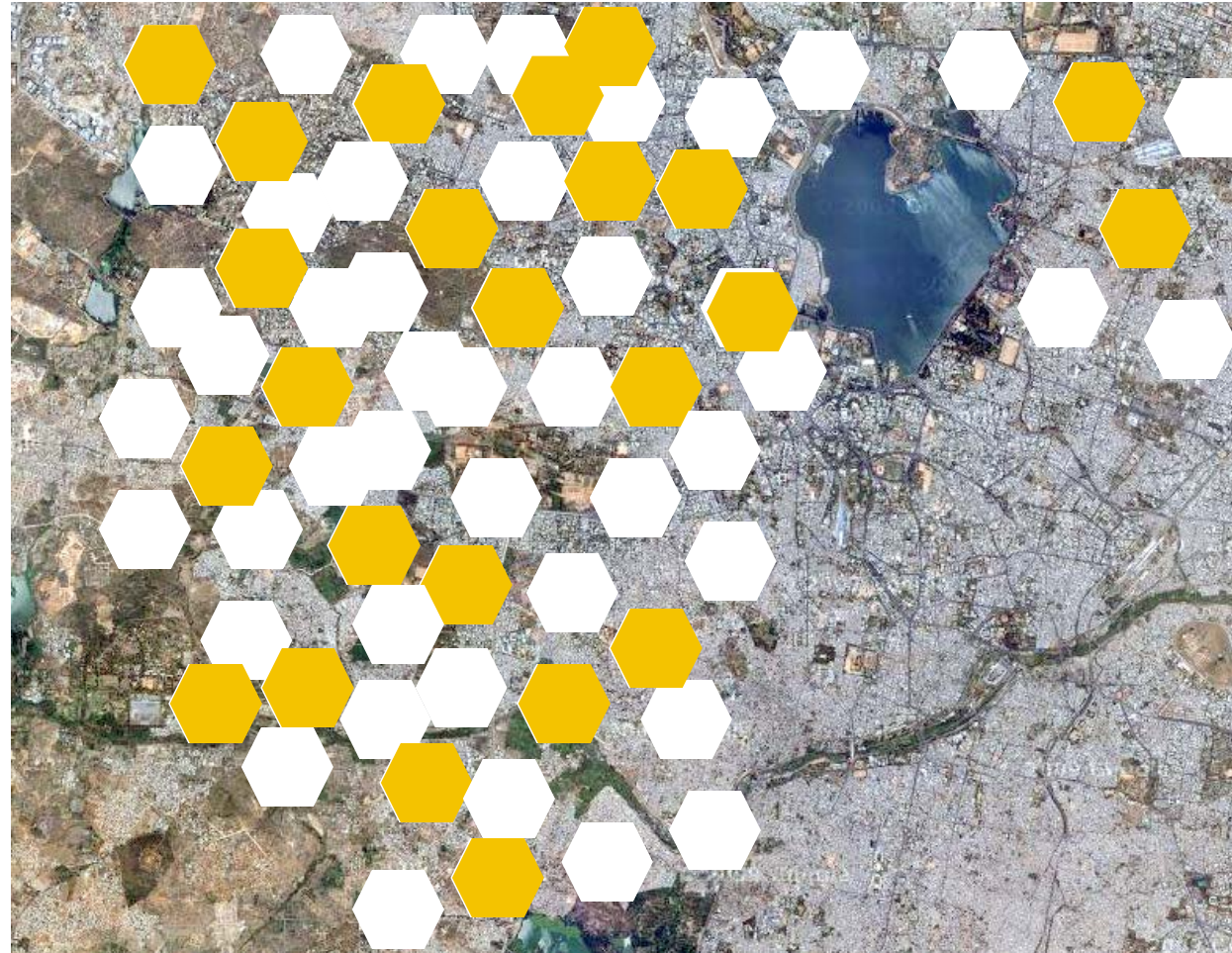
Purposive Assignment

Non-random treatment and control groups.



Random Sampling and Random Assignment

Randomly
sample
from area
of interest

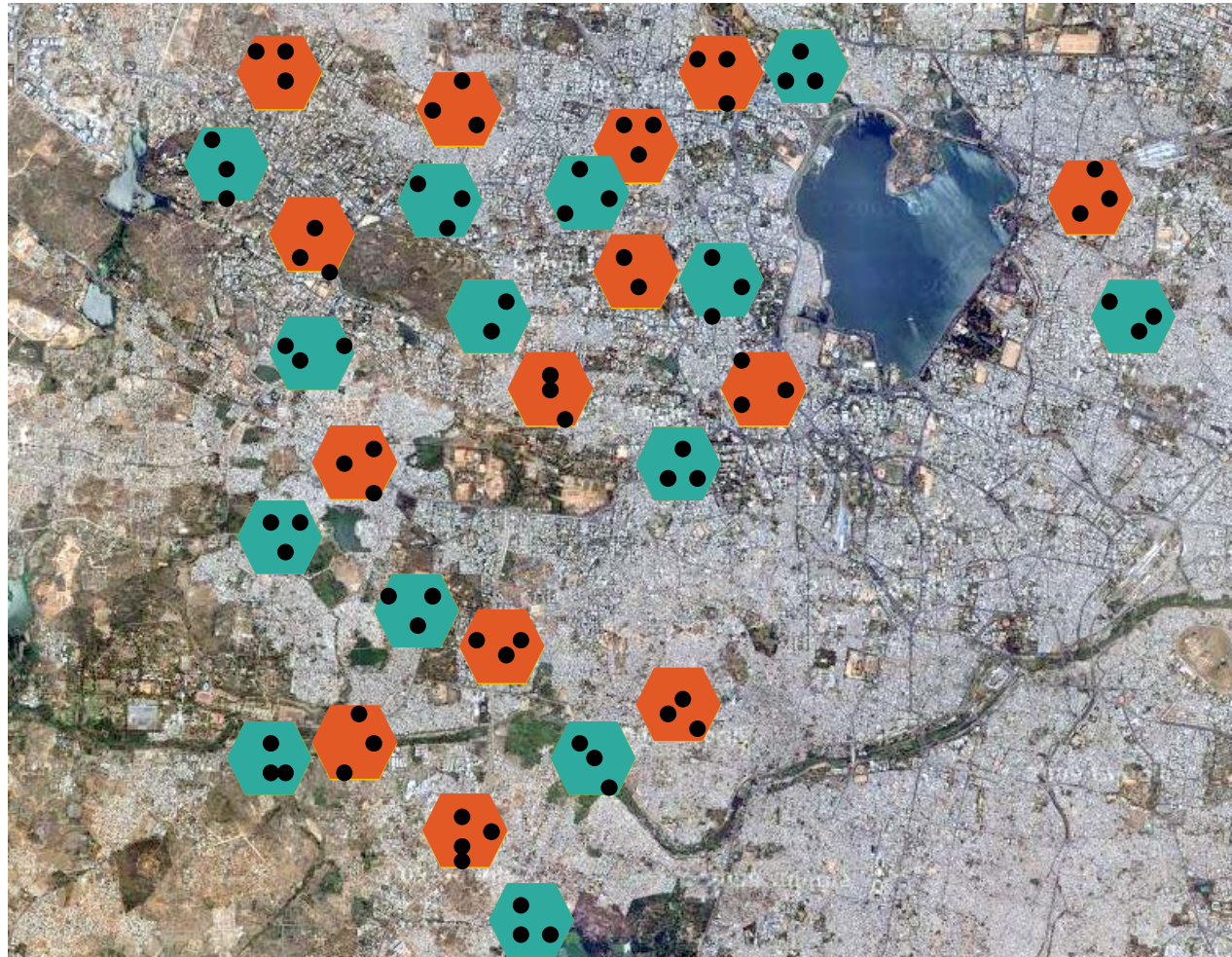


Random Sampling and Random Assignment

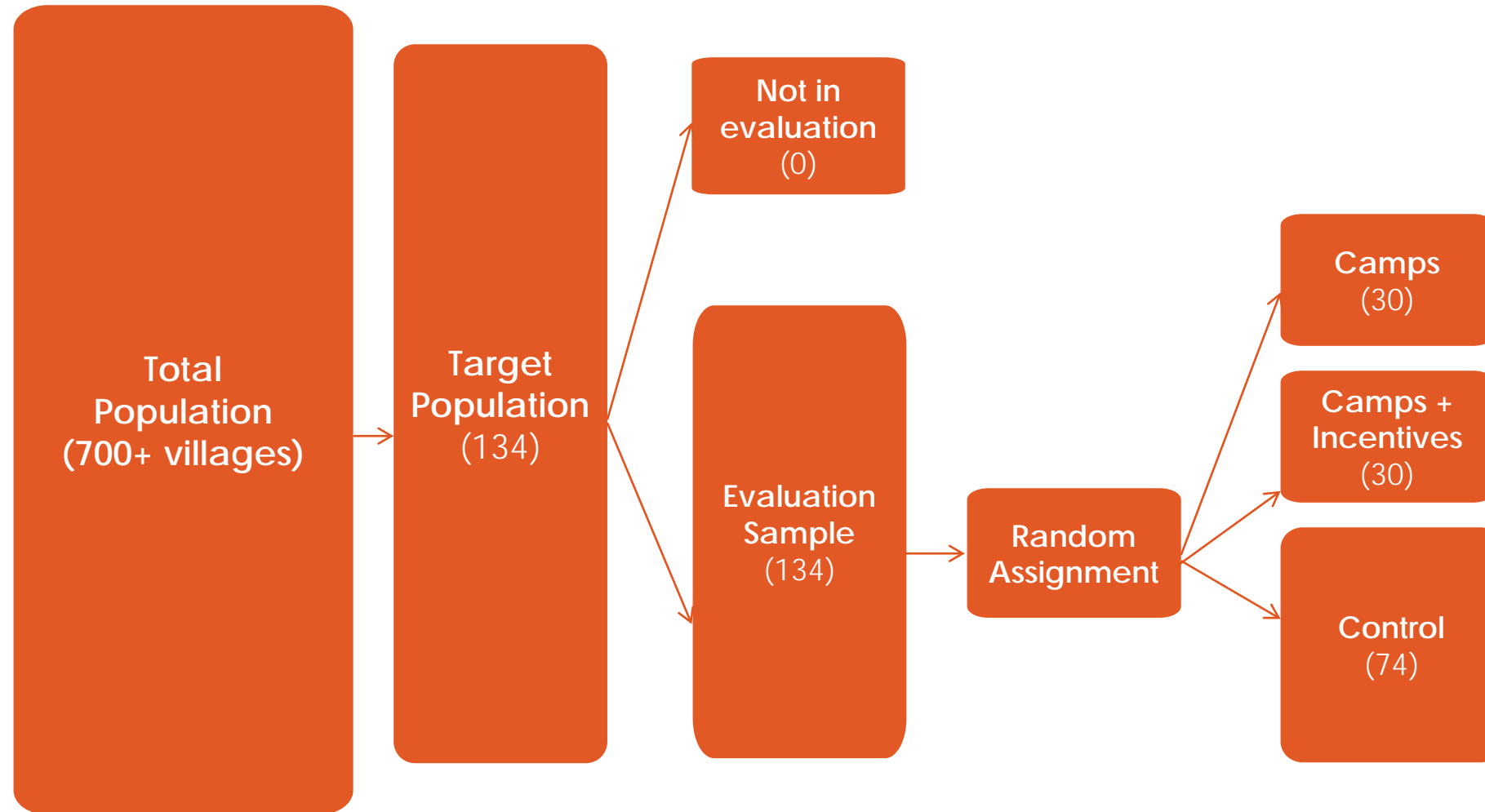
Randomly *sample* from area of interest

Randomly assign to **treatment** and **control**

Randomly *sample* from both treatment and control

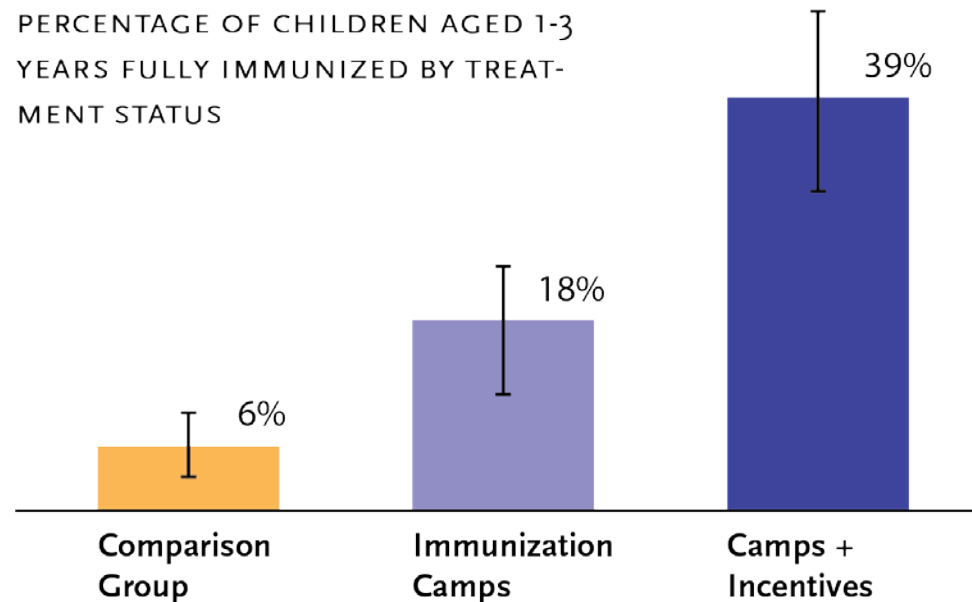


Immunization Example



Impact

- 6% full immunization in control villages
- 18% full immunization in camps villages
- 39% full immunization in camps + incentives villages



Designing Policy from Evidence

- Should we scale up this program nationally
 - How representative is rural Rajasthan? (recall: 22% vs. 44% nationally)
 - Do the same barriers to immunization exist in other parts of India?
 - What is the cost of this program to implement?
 - Do we need to replicate this study with a different version and/or in a different context?

Evidence-Based Policymaking

5. Cost-effectiveness Analysis



Incentives and Reliable Supply

FIGURE 2: PERCENTAGE OF CHILDREN AGED 1-3 YEARS FULLY IMMUNIZED BY TREATMENT STATUS

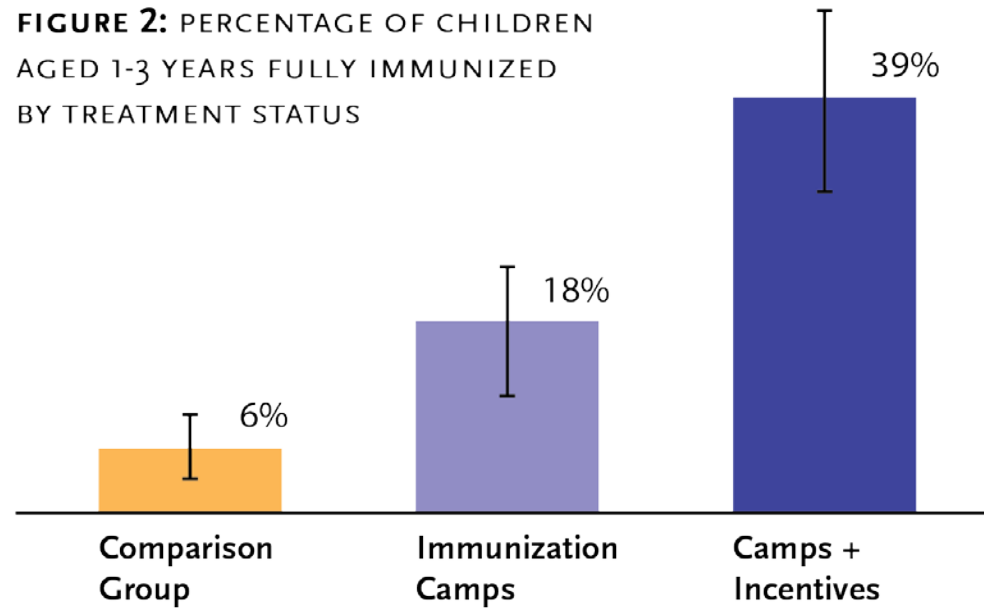
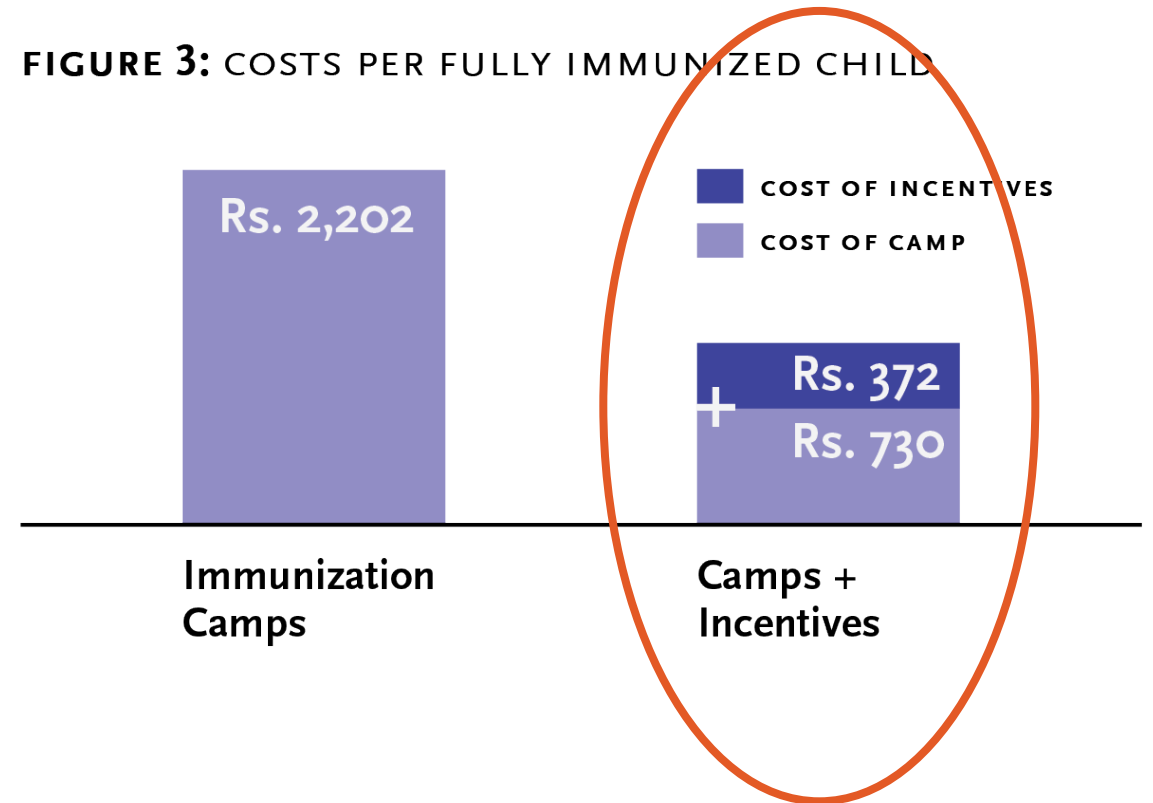
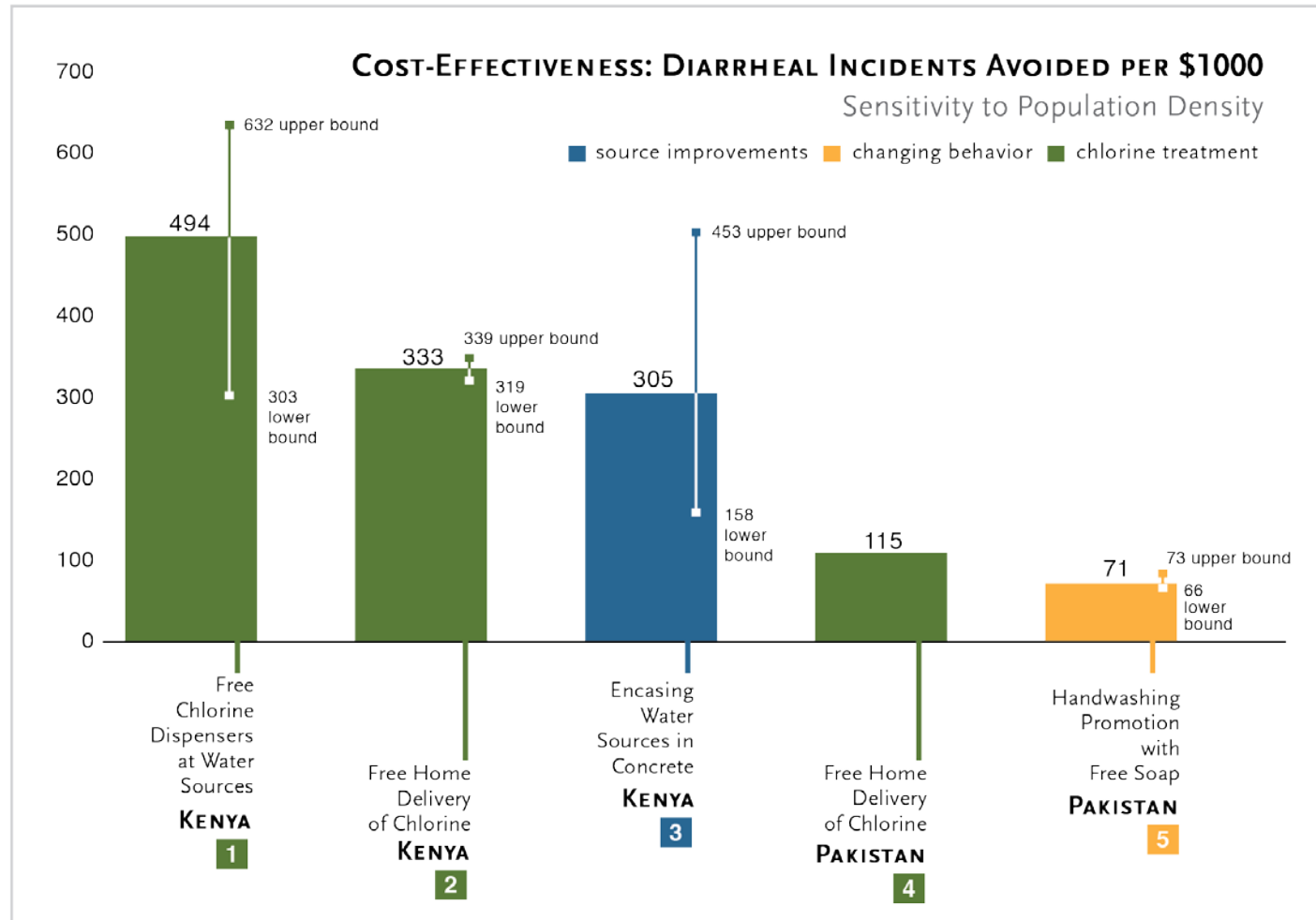


FIGURE 3: COSTS PER FULLY IMMUNIZED CHILD

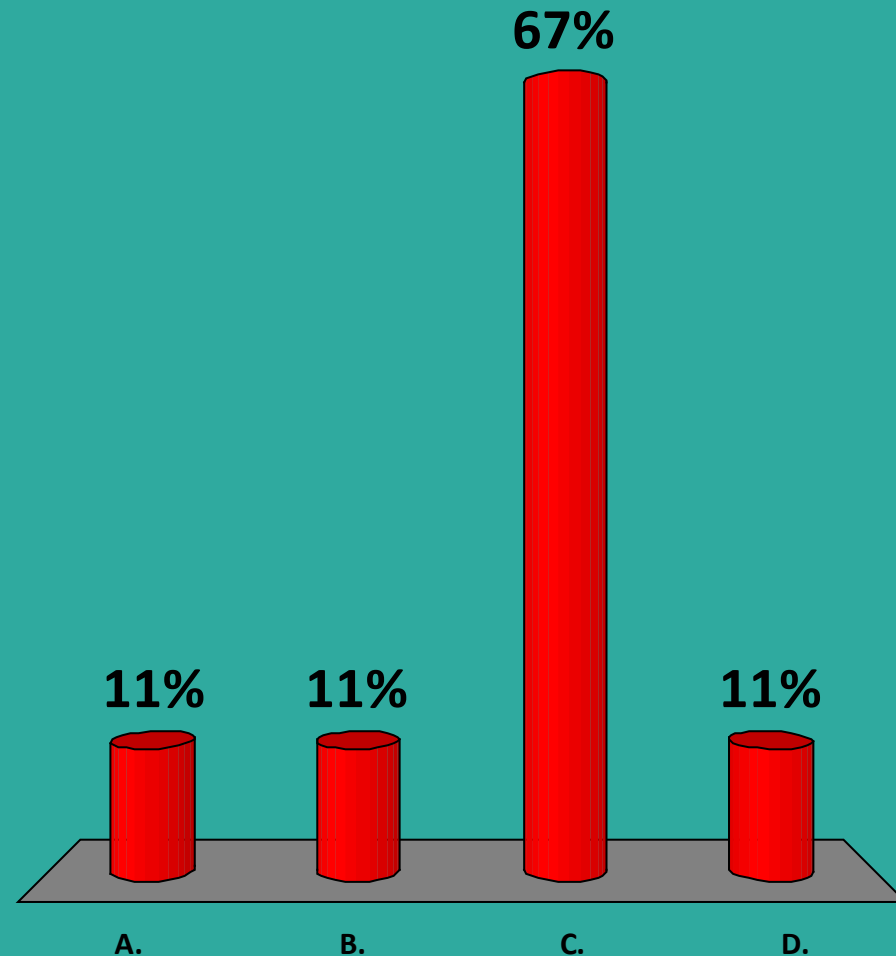


Cost-Effectiveness Diagram



When is a good time to do a randomized evaluation?

- A. After the program has begun and you are not expanding it elsewhere
- B. When a positive impact has been proven using rigorous methodology
- C. When you are rolling out a program with the intention of taking it to scale
- D. When a program is on a very small scale e.g one village with treatment and one without



When to do a randomized evaluation?

- When there is an important question for which you want/need to know the precise answer
- In terms of timing, not too early and not too late
- When the program and its implementation is representative and not gold plated (or tests a basic concept)
- You have the time, expertise, and money to do it right
- Develop an evaluation plan to prioritize

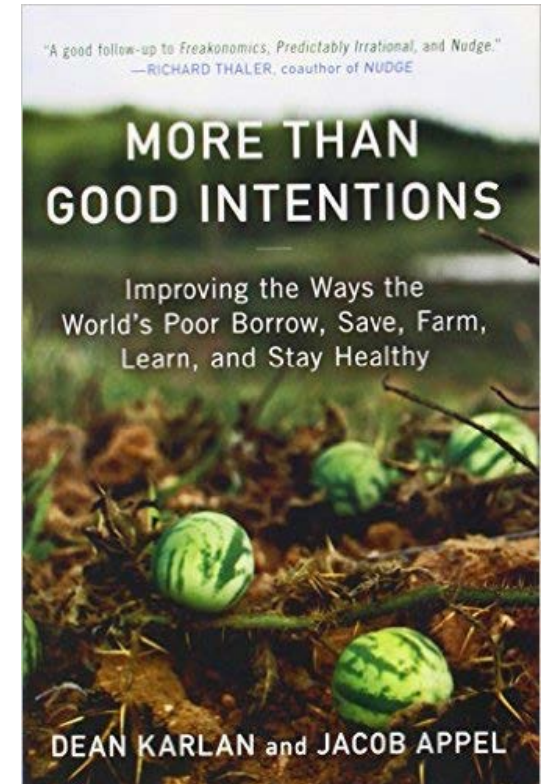
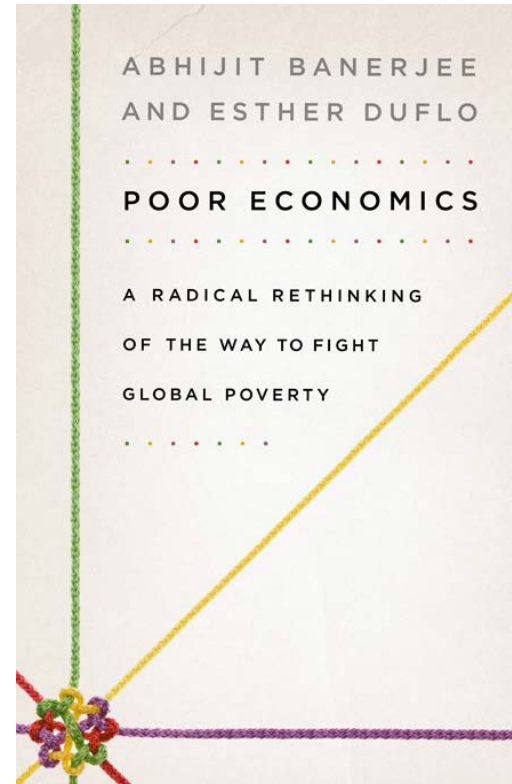
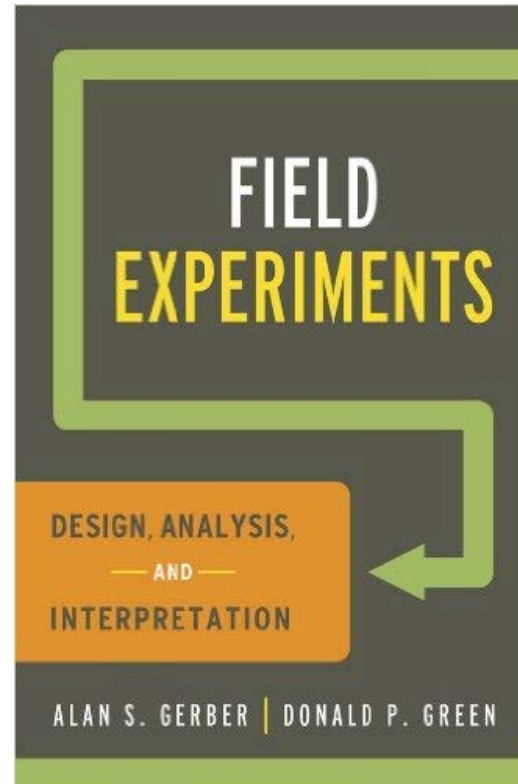
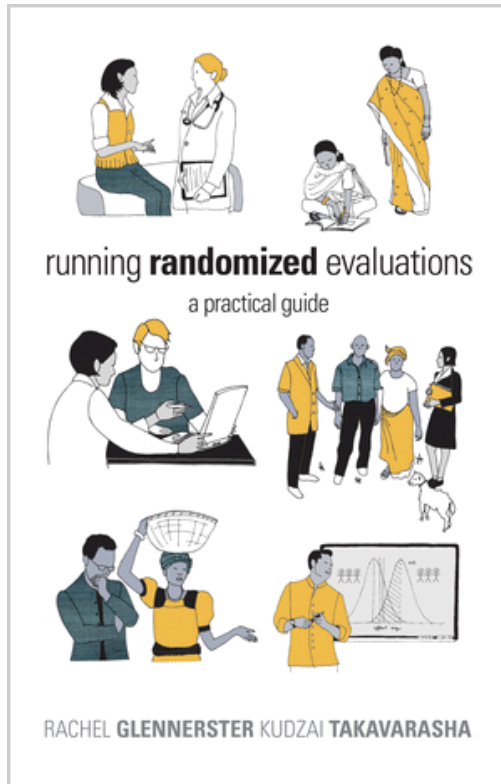
When **not** to do a randomized evaluation?

- When the program is premature and still requires considerable design work
- When the project is on too small in scale to randomize into at least two representative groups
- If a positive impact has been proven using rigorous methodology and resources are sufficient to cover everyone
- After the program has already begun and you are not expanding elsewhere

Developing an evaluation strategy

- Start with a question
- Verify the question hasn't been answered
- State a hypothesis
- Design the evaluation
- Determine whether the value of the answer is worth the cost of the evaluation
- With key questions answered from impact evaluations, process evaluation can give your overall impact
- A few high quality impact studies are worth more than many poor quality ones
- If you ask the right question, you're more likely to care

Some further readings



Thank you

