

# **Virtual Training Workshop on Counterfactual Impact Evaluation (CIE)**

## **Data Collection of Microdata in Difficult and Hard to Reach Areas**

**C4ED – EUTF**  
September 2022

## **Welcome to the Training Workshop on Counterfactual Impact Evaluation (CIE)**

The material of this workshop was produced with the financial support of the European Union. Its contents are the sole responsibility of C4ED and do not necessarily reflect the views of the European Union

# Introductions

# Communication during the training

---



**MUTE BUTTON**



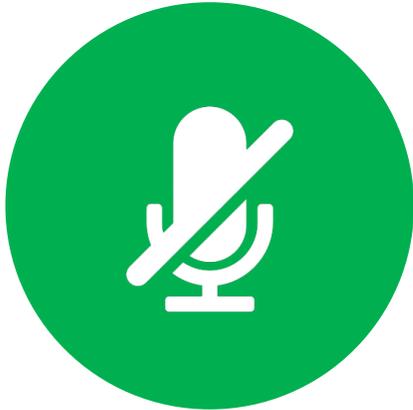
**QUESTIONS**



**FEEDBACK**

# Communication during the training

---



**MUTE BUTTON**



**QUESTIONS**



**FEEDBACK**

# Communication during the training

---



**MUTE BUTTON**



**QUESTIONS**



**FEEDBACK**

# Asking Questions

---

- Please post your questions in the chat room
- Like questions of others, so we know they are particularly relevant or urgent



- Carolin will read out all questions and we will answer these at once
- Use the longer breaks to ask more questions

# Communication during the training

---



**MUTE BUTTON**



**QUESTIONS**



**FEEDBACK**

# Asking Questions

---

- Please make suggestions
- Feel free to share your comments
- More feedback and questions (especially for the Q&A session):

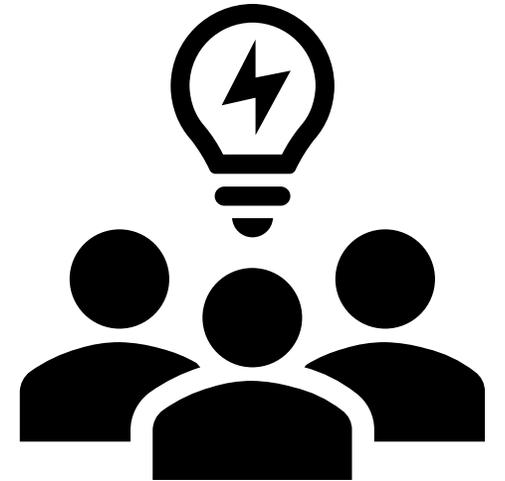
# Day 1 Agenda

Local Time	Proposed activities
10:00 – 10:30 (30 min)	Welcome words by EUTF and C4ED
10:30 – 11:15 (45 min)	Session 1: Recap of Counterfactual Impact Evaluation (CIE) Methods
11:15 – 11:40 (20 min)	Session 2: Overview of data collection and its relevance in CIE and Monitoring Systems (MS).
11.40 – 12.10 (30 min)	Break
12.10 – 12.30 (20 min)	Breakout Session 1
12:30 – 13:10 (40 min)	Session 3a: Stages of Preparation for Data Collection in CIE - Moving from Evaluation Questions to Indicators
13.10 – 13.55 (45 min)	Lunch Break
13:55 – 14:25 (30 min)	Session 3b: Stages of Preparation for Data Collection in CIE - Developing Data Collection Tools
14:25 – 15:05 (40 min)	Session 4a: Sampling in data collection: Challenges and implications - Developing a Sample Frame.
15:05 – 15:30 (25 min)	Session 4b: Sampling in data collection: Challenges and implications - Taking a Sample
15.30 – 15.45 (15 min)	Breakout Session 2
15:45 – 16:00 (15 min)	End of Day Quiz and Q&A

# Breakout Sessions

---

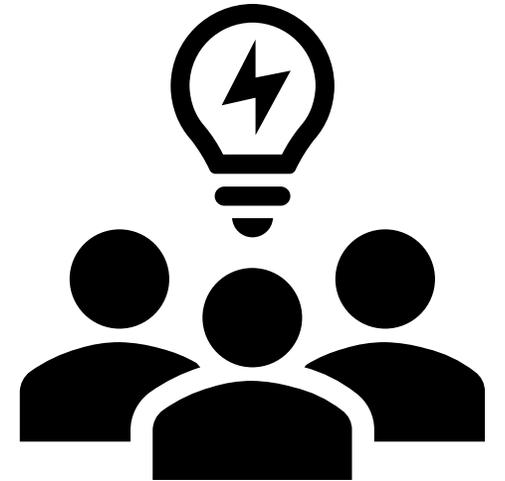
- During the workshop we will have breakout sessions to allow discussion on the content.
- The breakout sessions will take the form of a group task and be conducted using Microsoft Teams and managed by Nathan.
- Participants will be organized into groups based on the EUTF project – using the response to the initial poll.



# Breakout Sessions

---

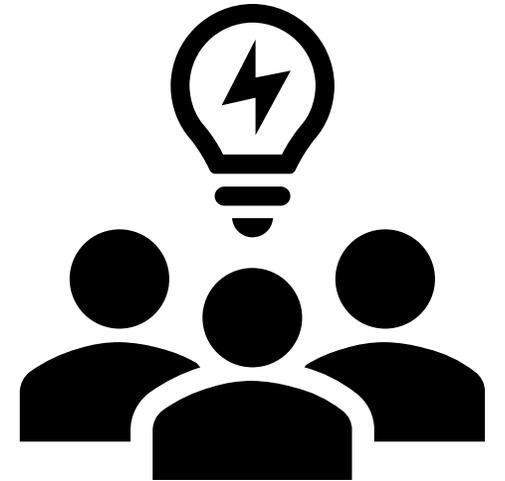
- Anyone outside of a project/country will be assigned to a team at random.
- If a group of more than 5 people are sat together they can form their own breakout group – also using the response to the initial poll.



# Breakout Sessions

---

- During the breakout rooms, Eli and Nathan will drop in and out of rooms and be available for questions, provide support and to review progress



# **Session 1: Recap of CIE Methods**

**C4ED – EUTF**  
September 2022

# Session 1 Learning Objectives

---

- Participants will be able to describe what is meant by the term Counterfactual Impact Evaluation
- Participants will be able explain that characteristics of a good counterfactual for a CIE
- Participants will be able explain the basics ideas behind four common approaches to simulating the counterfactual
- Participants will be directed to where non-technical reference material can be found if they wish to dig deeper into CIE methods

# Refresh knowledge on CIE

---

- In the first part of the session we will refresh our knowledge on Counterfactual Impact Evaluation (CIE)
- We will briefly review the basics on counterfactuals, common methods of **identifying impact** and the benefits of using CIE
- For a more in-depth discussion of CIE, please review the slides from last years training: *Training Workshop on Counterfactual Impact Evaluation (CIE)*
- We will also share useful external resources and case studies on CIE
- <https://europa.eu/capacity4dev/>

# Refresh knowledge on CIE

---

- We'll start with a short quiz on CIE Methods
- Kindly look in the chat section for the link to the quiz
- Enter your name and use the code below if required
  - Code: **35890009**

**What is a counterfactual impact evaluation?**

# Refresh knowledge on CIE

---

- **What is a counterfactual impact **evaluation**?**
- Evaluation is the systematic assessment of the design, implementation or results of an initiative for the purposes of learning or decision-making.

# Refresh knowledge on CIE

---

- **What is a counterfactual **impact** evaluation?**
- Impact should be considered as the causal effect of the program/policy
- This is the effect on outcomes of interest that the program/policy directly causes and can be directly attributable to the program

# Refresh knowledge on CIE

---

- What is a **counterfactual** impact evaluation?
- The counterfactual is the outcome, measured at that same point in time, had the program had not been introduced to the beneficiaries.
- Fundamental problem: it is impossible to measure or observe the counterfactual
- Solution: Simulate counterfactual situation
  - creating a control/comparison group

## How is a CIE designed?

**Goal:** Simulate counterfactual situation with a comparison group

- The comparison group:
  - Has the same characteristics (on average) as the treatment group
  - Is not exposed to the program
  - Would react similarly to the program as the treatment group (if it were to participate)
- Based on the intervention design and context, timeline, data availability and budget, the most appropriate approach to use is selected:
  - Experimental methods
  - Quasi-experimental methods

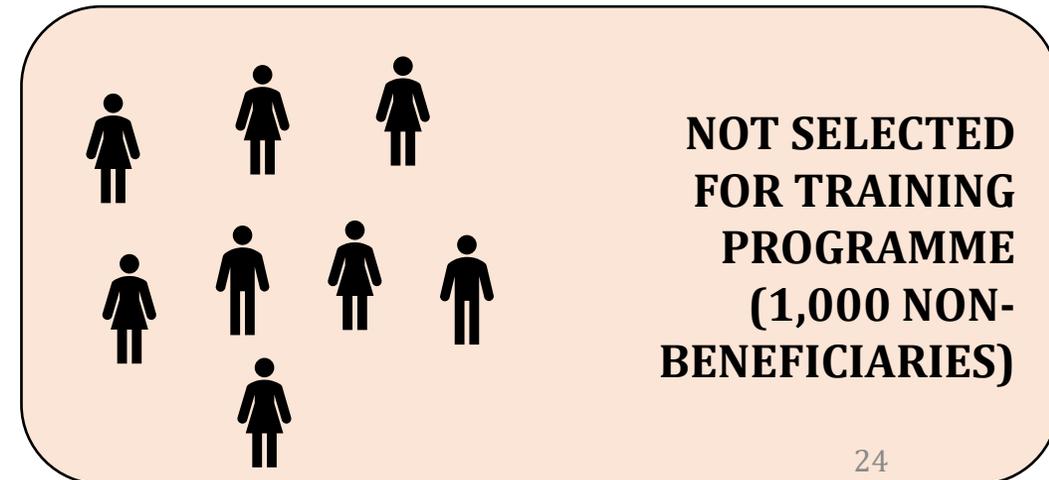
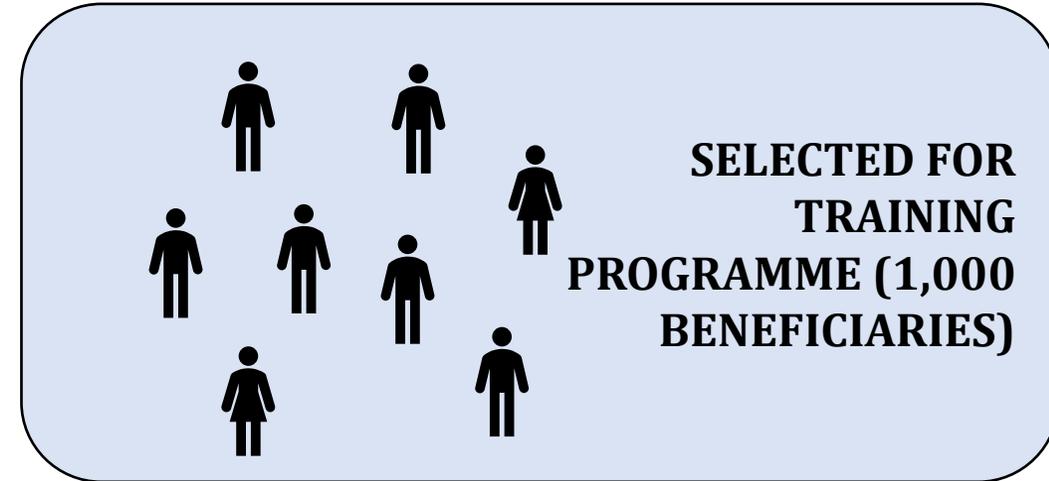
# Simulating a counterfactual

---

- In the following example we consider the selection for a youth vocational training program aimed to improve employment outcomes.
- In the first case, applicants are invited to take part in short interviews to discuss their application.
- Based on their application and interview, applicants are either selected to take part in a vocational training programme, or not.

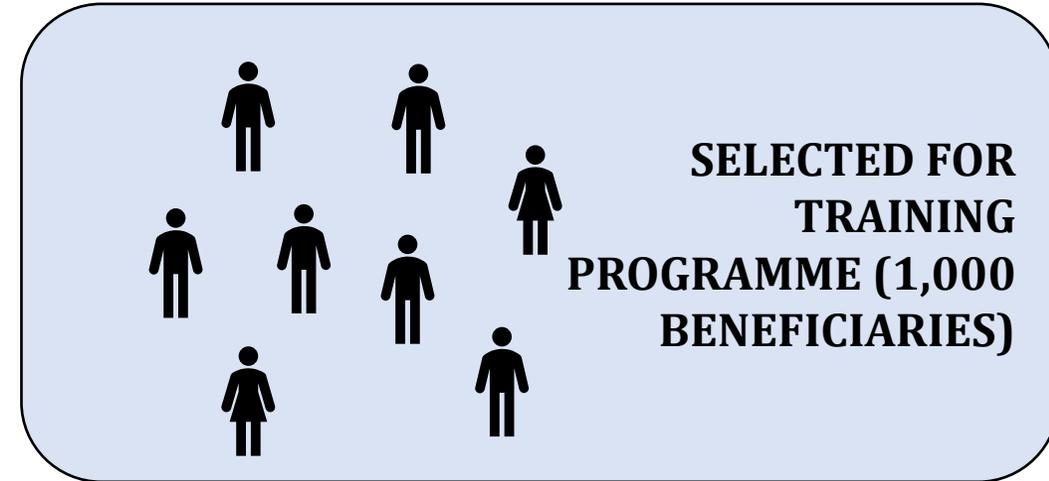
# Applicants interview for a place

Number of Applicants: 2,000 (1,000 Males / 1,000 Females)



# Applicants interview for a place

Number of Applicants: 2,000 (1,000 Males / 1,000 Females)



Would this group make a  
good counterfactual?



# Sources of difference between groups

---

- Applicants that are selected are likely to be quite different from those that are not as 'stronger' applicants are chosen.
- For a program manager this may be desirable to have the most able in the program
- However – for an evaluator looking to measure the impact of the program, these differences will mean that the groups are not easily comparable
- **As it stands they would not represent a good counterfactual**

# Group comparison



Variable	Average
Age	30
Years of schooling	10
Previous employment	60%
Parent income	5,000



Variable	Average
Age	24
Years of schooling	6
Previous employment	46%
Parent income	3,400

# Group comparison



Motivation



Self-confidence



Determination



Motivation



Self-confidence



Determination



# Counterfactual Selection

---

- In the previous example – simply using all rejected applicants is a poor counterfactual
- How could you design the process differently to be able to use rejected applicants as a counterfactual?



## **Randomized Controlled Trial**

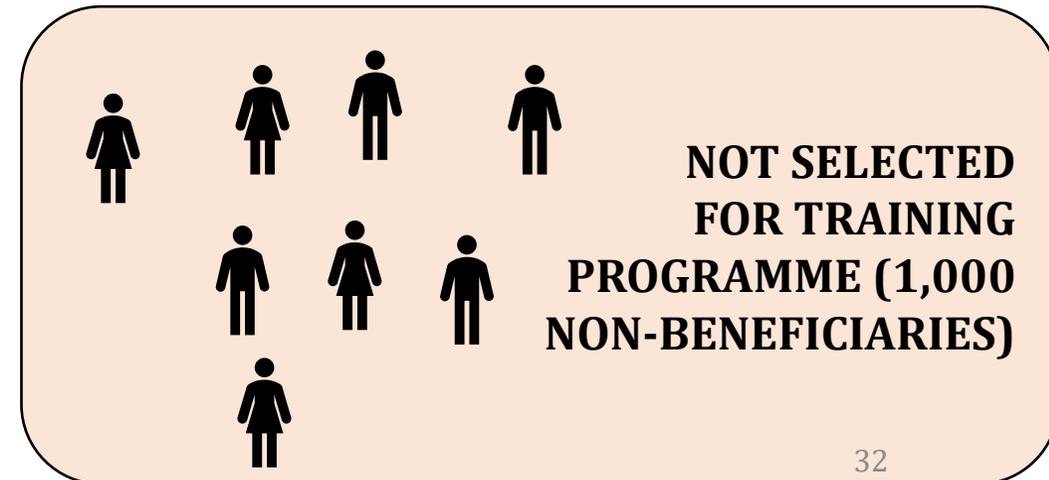
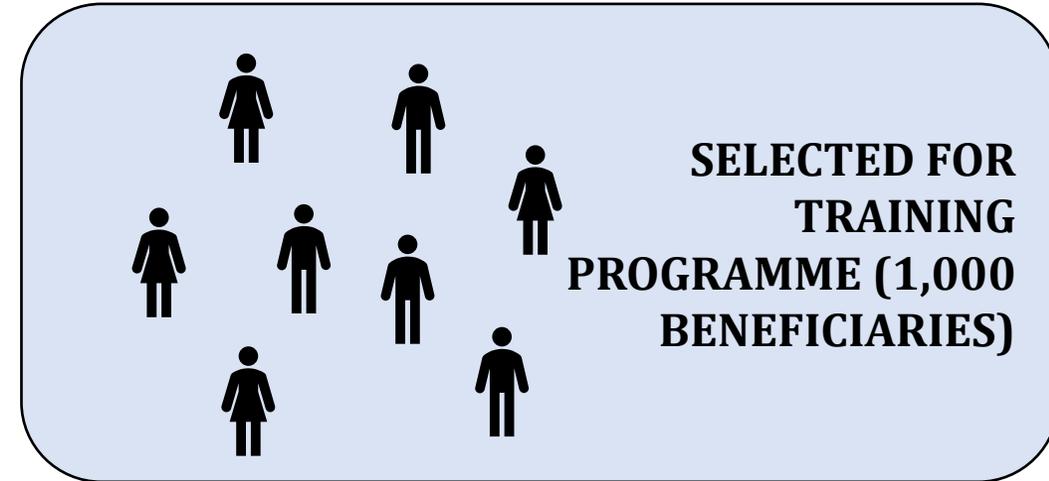
# Randomized Assignment

---

- If you integrated an evaluation into the program implementation, you could create a counterfactual through using a lottery to decide who is selected
- This is called **randomized assignment**
- Why does this help?
- Assuming the number of applicants is large enough, the two randomly assigned groups will be similar

# Randomized selection

Number of Applicants: 2,000 (1,000 Males / 1,000 Females)



# Group comparison: Randomized



Variable	Average
Age	27
Years of schooling	8
Previous Employment	52%
Parent Income	4,300



Variable	Average
Age	27
Years of schooling	8
Previous Employment	54%
Parent Income	4,100

# Group comparison: Randomized



Motivation



Self-confidence



Determination



Motivation



Self-confidence

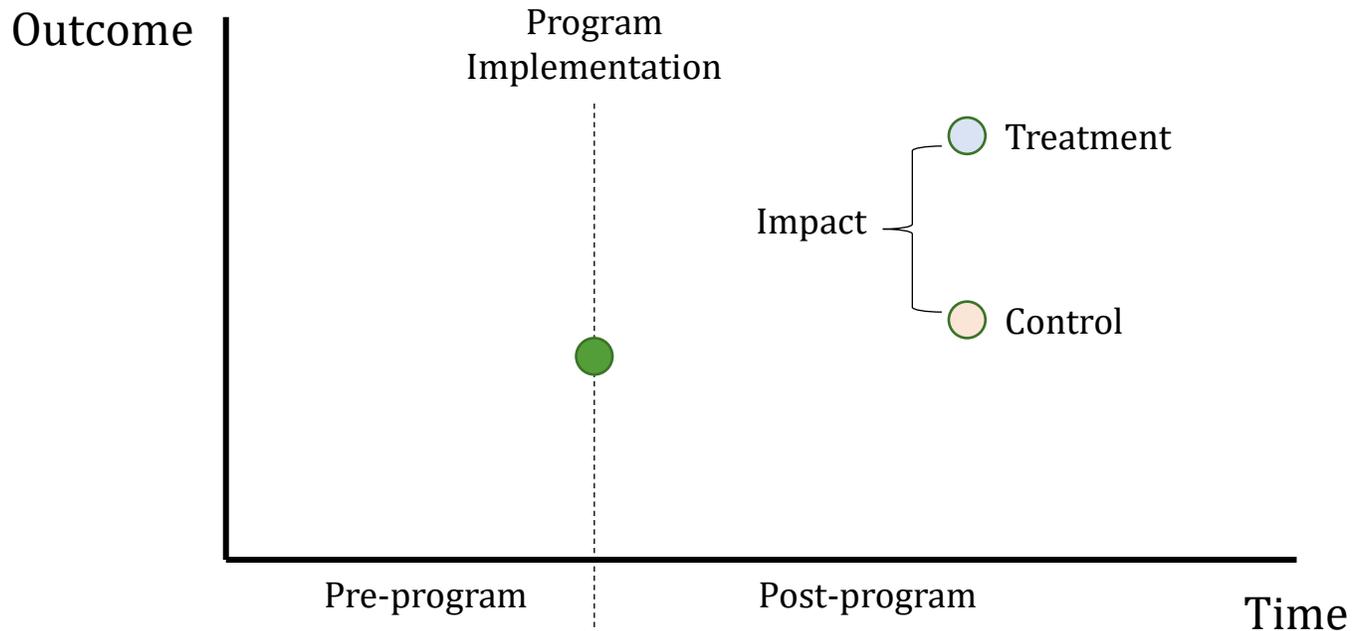


Determination



# Measuring Impact in an RCT

- Because randomized assignment creates two groups that are (on average) comparable at the beginning of the programme. The impact can be measured simply as the difference in the outcome after the program.



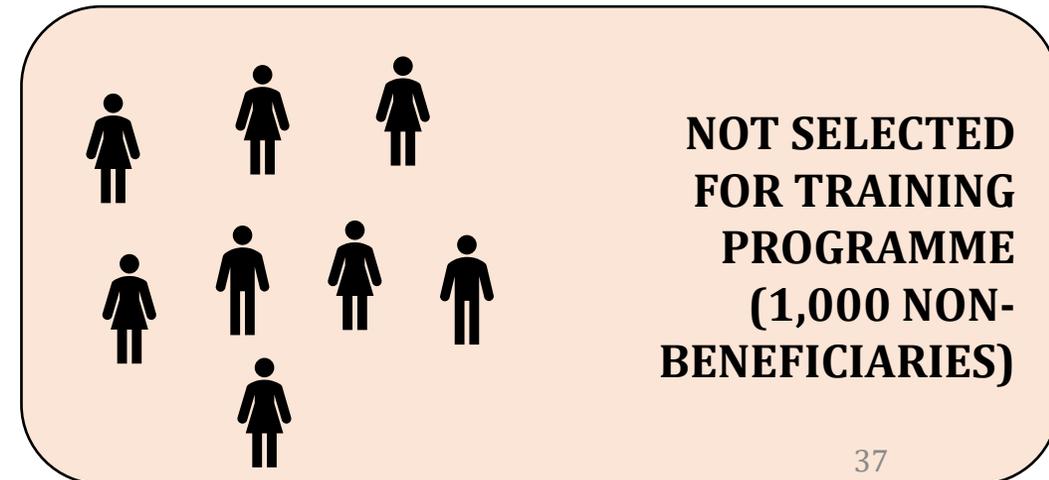
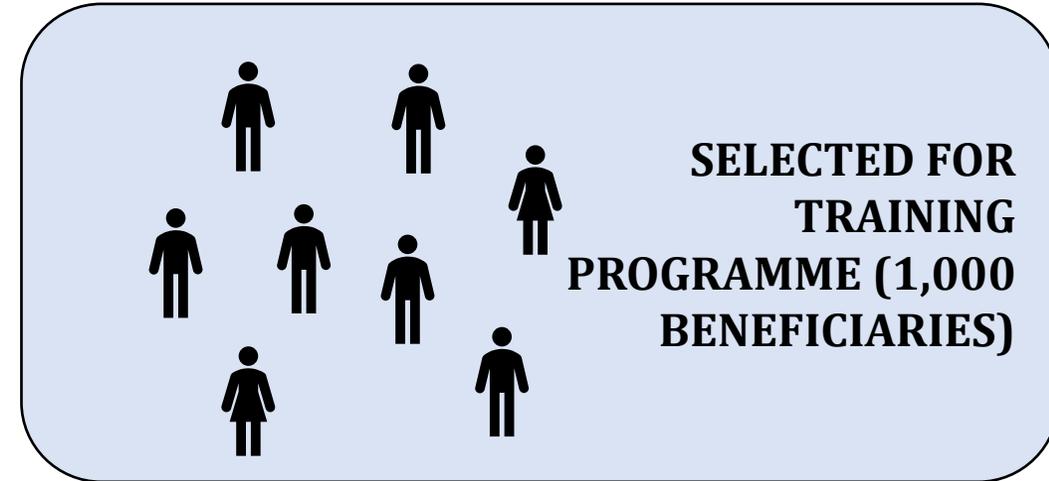
# Quasi-experimental methods

---

- If randomized selection is not possible:
  - Randomization may not be socially or politically acceptable
  - Randomization may not be feasible
  - CIE is designed only after implementation starts
- May be possible to use quasi-experimental methods to construct counterfactual.
- In the following slides we will go through the basic idea of using matching and difference-in-difference approaches to simulating a counterfactual.

# Applicants interview for a place

Number of Applicants: 2,000 (1,000 Males / 1,000 Females)



# Group comparison



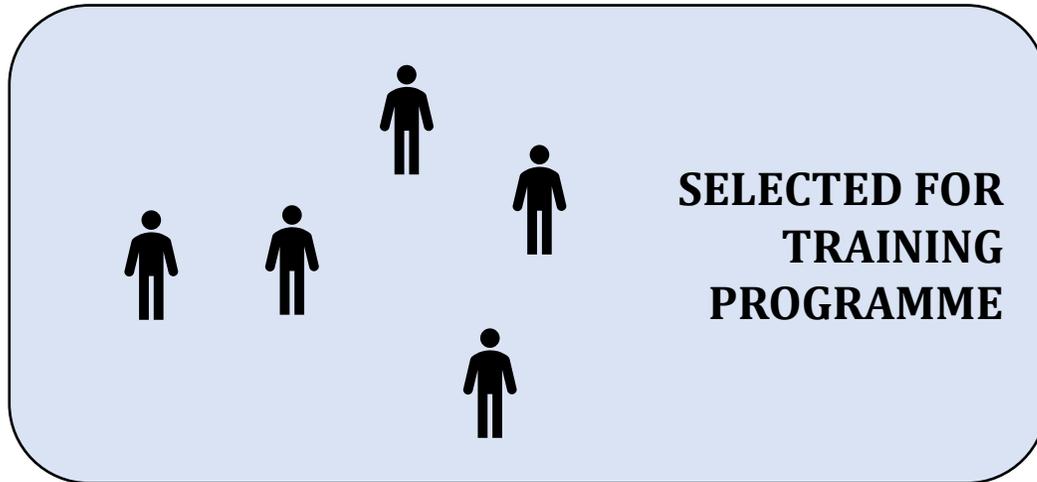
Variable	Average
Age	30
Years of schooling	10
Previous employment	60%
Parent income	5,000



Variable	Average
Age	24
Years of schooling	6
Previous employment	46%
Parent income	3,400

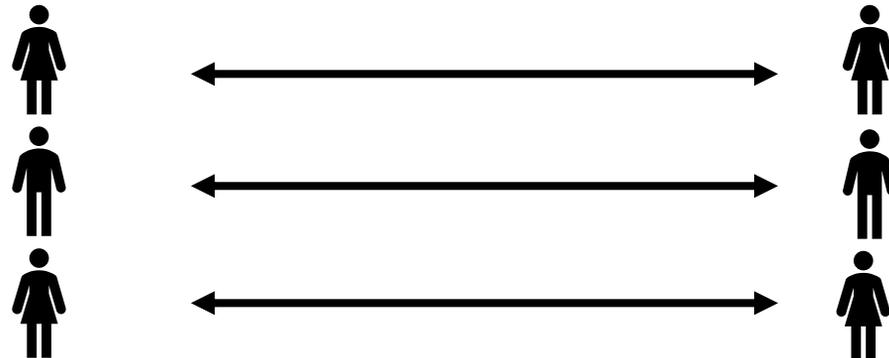
## Matching

# Matching

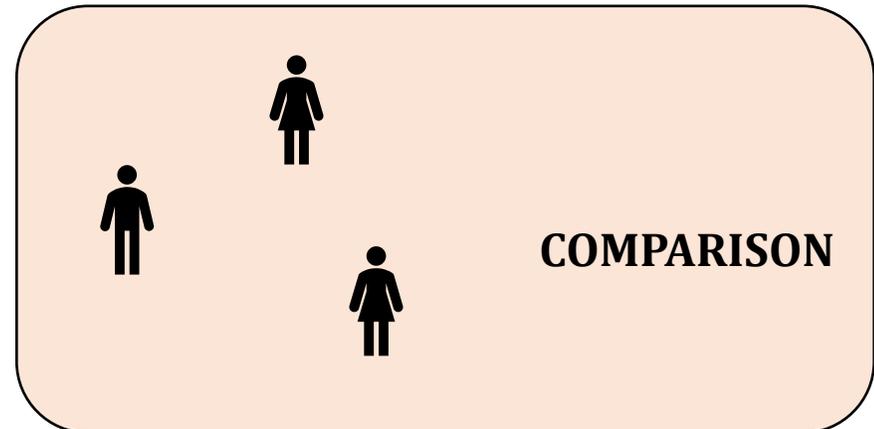
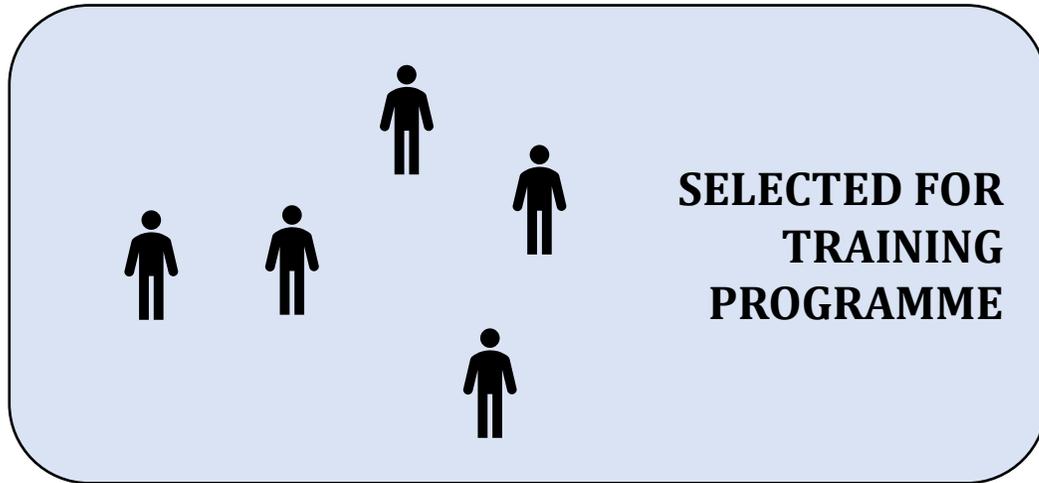


Look for pairs of people that are similar in terms of characteristics

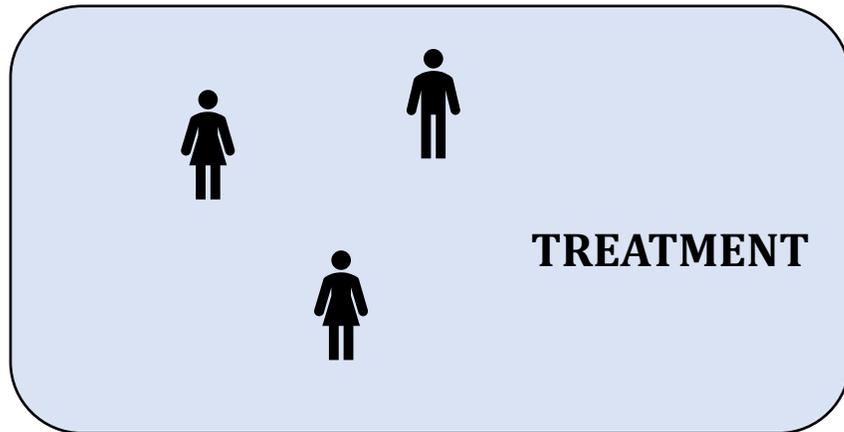
(Before program or do not change over time)



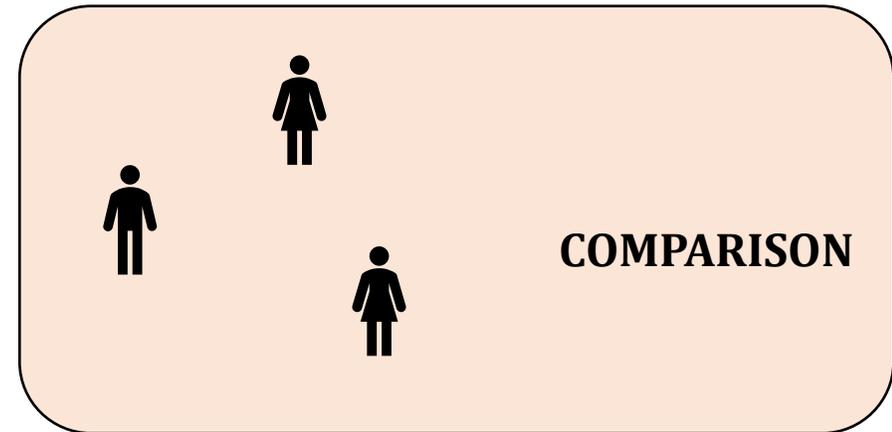
# Matching



# Matching



Variable	Average
Age	28
Years of schooling	8
Previous employment	54%
Parent income	4,200

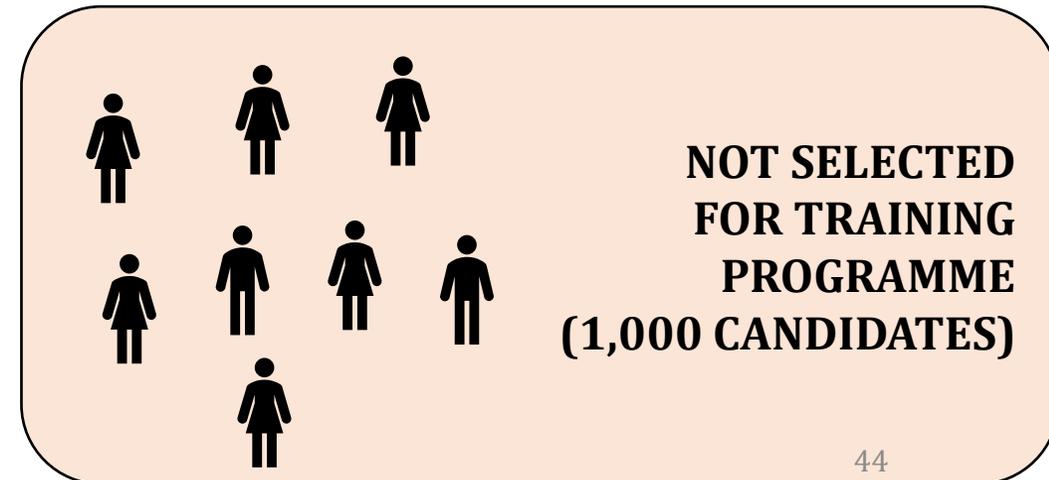
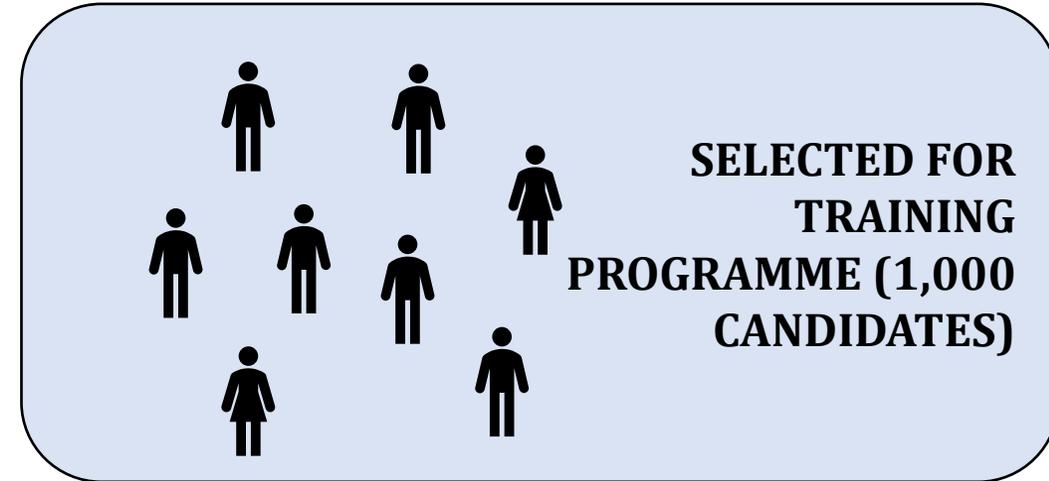


Variable	Average
Age	27
Years of schooling	7
Previous employment	53%
Parent income	4,300

## **Difference-in-differences**

# Applicants interview for a place

Number of Applicants: 2,000 (1,000 Males / 1,000 Females)



# Group comparison

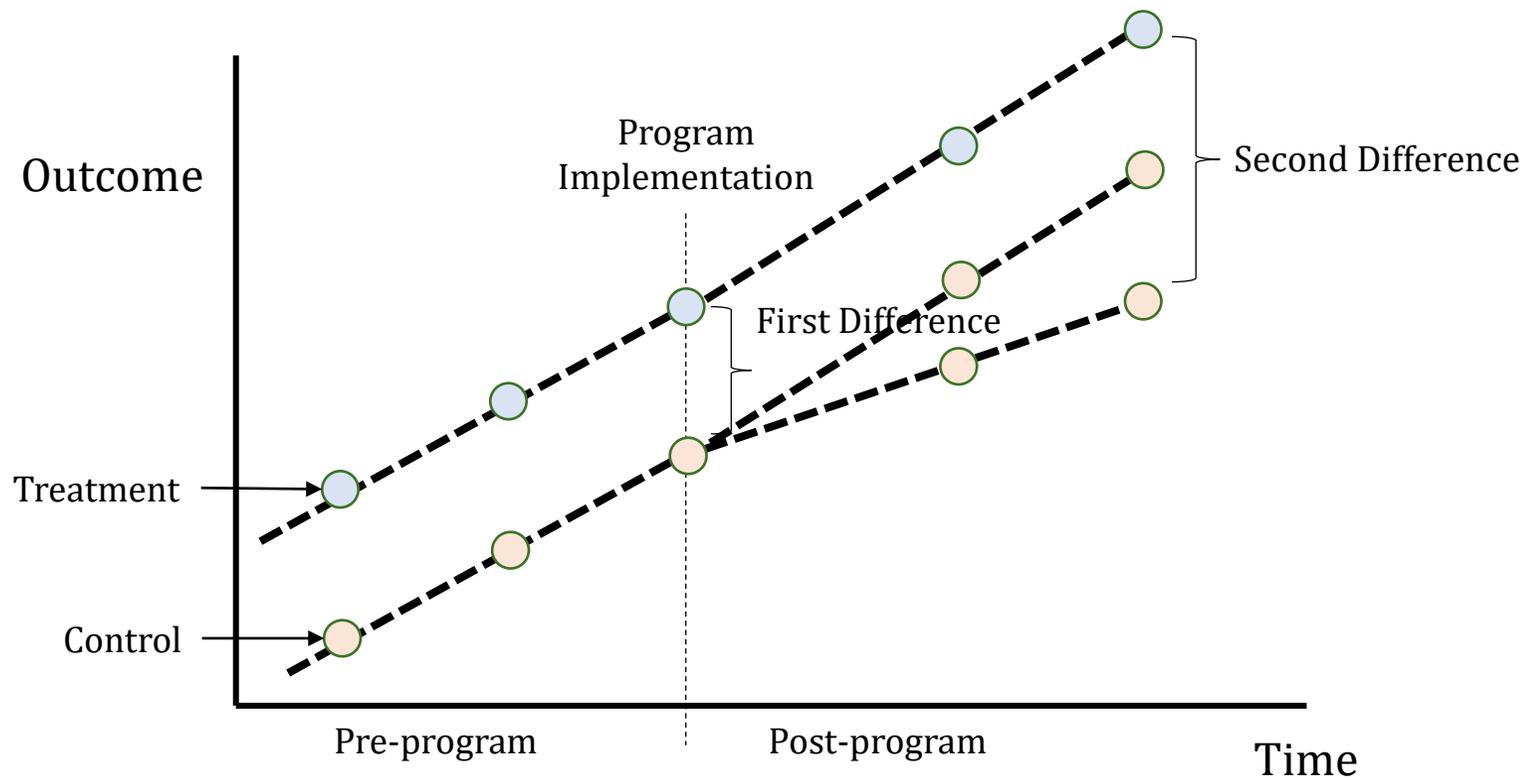


Variable	Average
Age	30
Years of schooling	10
Previous Employment	60%
Parent Income	5,000

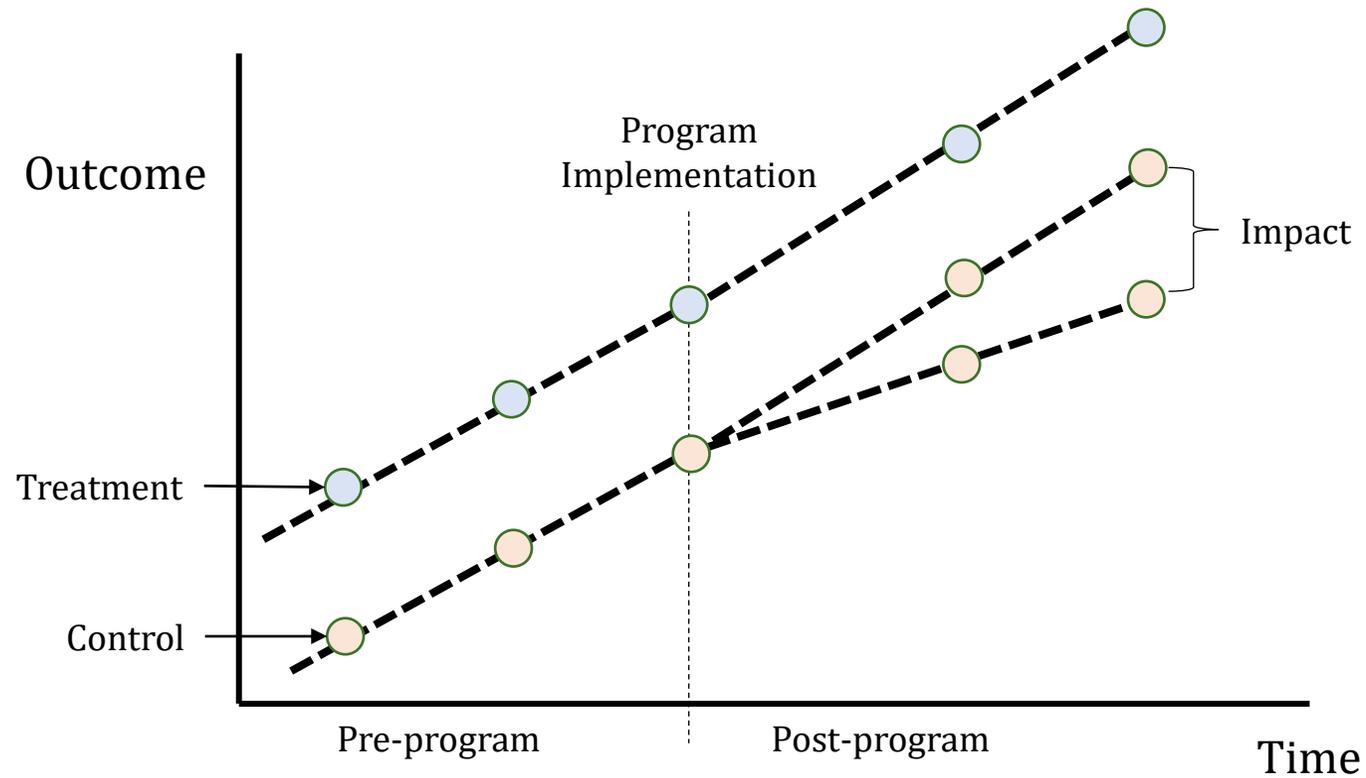


Variable	Average
Age	24
Years of schooling	6
Previous Employment	46%
Parent Income	3,400

# Measuring Impact in an RCT



# Measuring Impact in an RCT



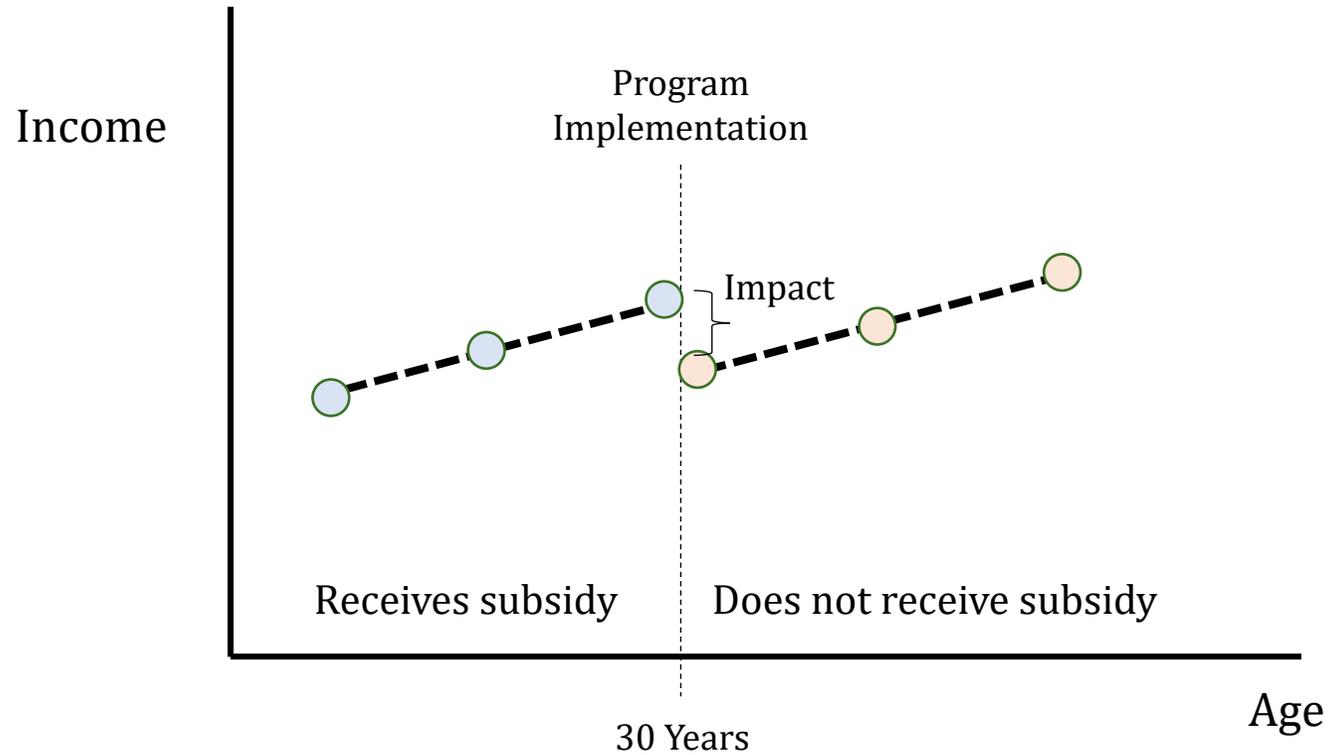
## **Regression-Discontinuity Design**

# Regression-Discontinuity Design

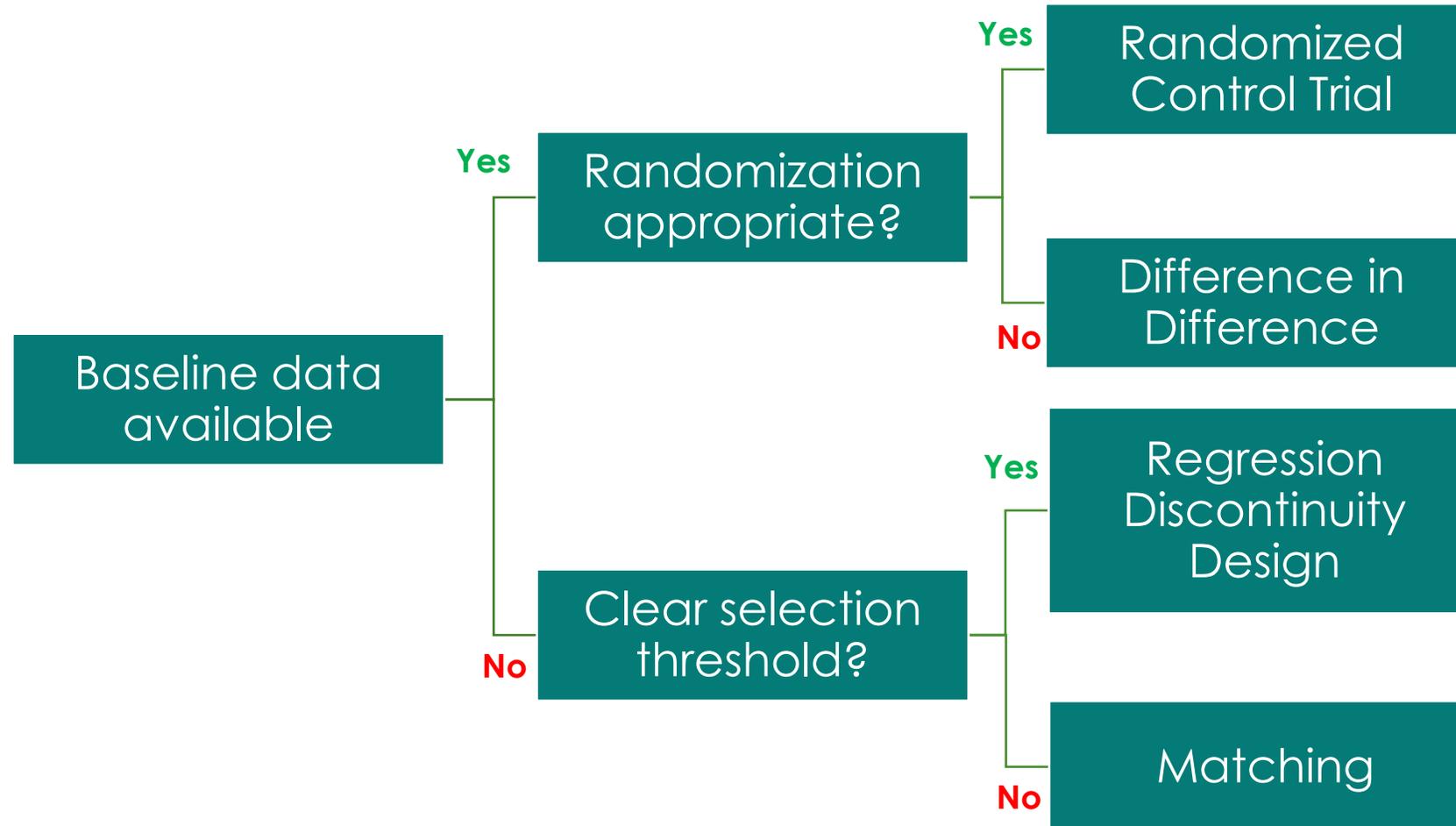
---

- In cases where there is a clear rule assigning individuals to take part in a program or not,
- We know who just missed out on participation and those who were marginally selected
- An example of this would be a national youth wage subsidy for unemployed people below the age of 30

# Measuring Impact in an RCT



# Which Approach to take



- **Why do a CIE?**
- To determine whether an intervention creates an **attributable, causal change** in the outcome, to what **magnitude and how** (the **causal mechanism**)
- To **learn** which intervention strategy works best
- To help **inform** policymakers to make evidence-based **decisions**

- Training Workshop on Counterfactual Impact Evaluation (CIE) – PowerPoint Slides

## Books

- [World Bank, Impact Evaluation in Practice - Second Edition \(Book\)](#)

## Videos

- [InterAction, Introduction to Impact Evaluation](#)
- [Esther Duflo, Randomized Controlled Trials and Policy Making in Developing Countries](#)

## Podcasts

- [IEU Talks Episode 2: The Power of Impact Evaluation in Development Cooperation](#)
- [Evidencing impact \(parts 1+2\)](#)

# END OF SESSION 1

## **Session 2: Overview of data collection and its relevance in Counterfactual Impact Evaluations and Monitoring Systems.**

**C4ED – EUTF**  
September 2022

## Session 2 Learning Objectives

---

- Participants will be able to explain why high quality data is necessary for an impact evaluation
- Participants will be able to describe the difference between quantitative and qualitative data
- Participants will be able to apply their knowledge to provide examples of data sources to assess the Evaluation Questions and Theory of Change of a program.

# Data Collection in CIE

---

- In this workshop we will focus on collecting micro-data in hard-to-reach areas
- Hard-to-reach areas can mean various things:
  - Geographic - remote
  - Insecurity and fragile contexts
  - “Invisible” Populations

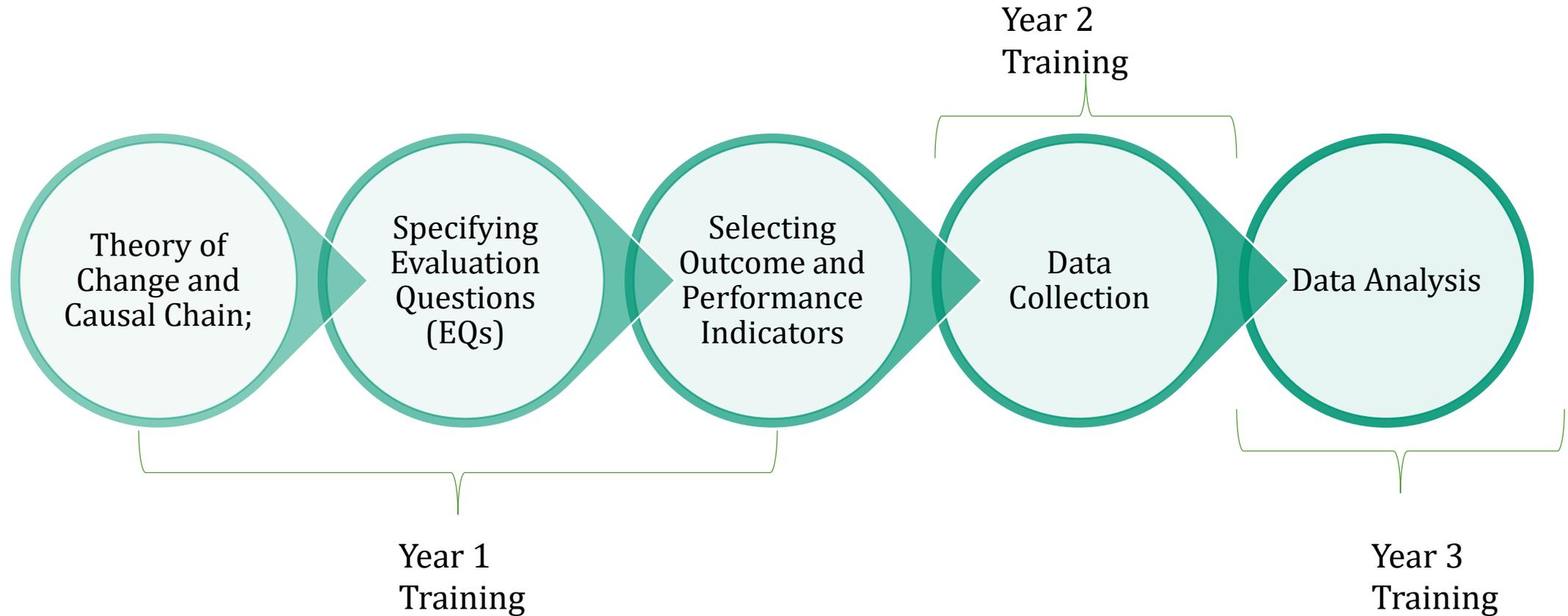
# Data Collection and CIE

---

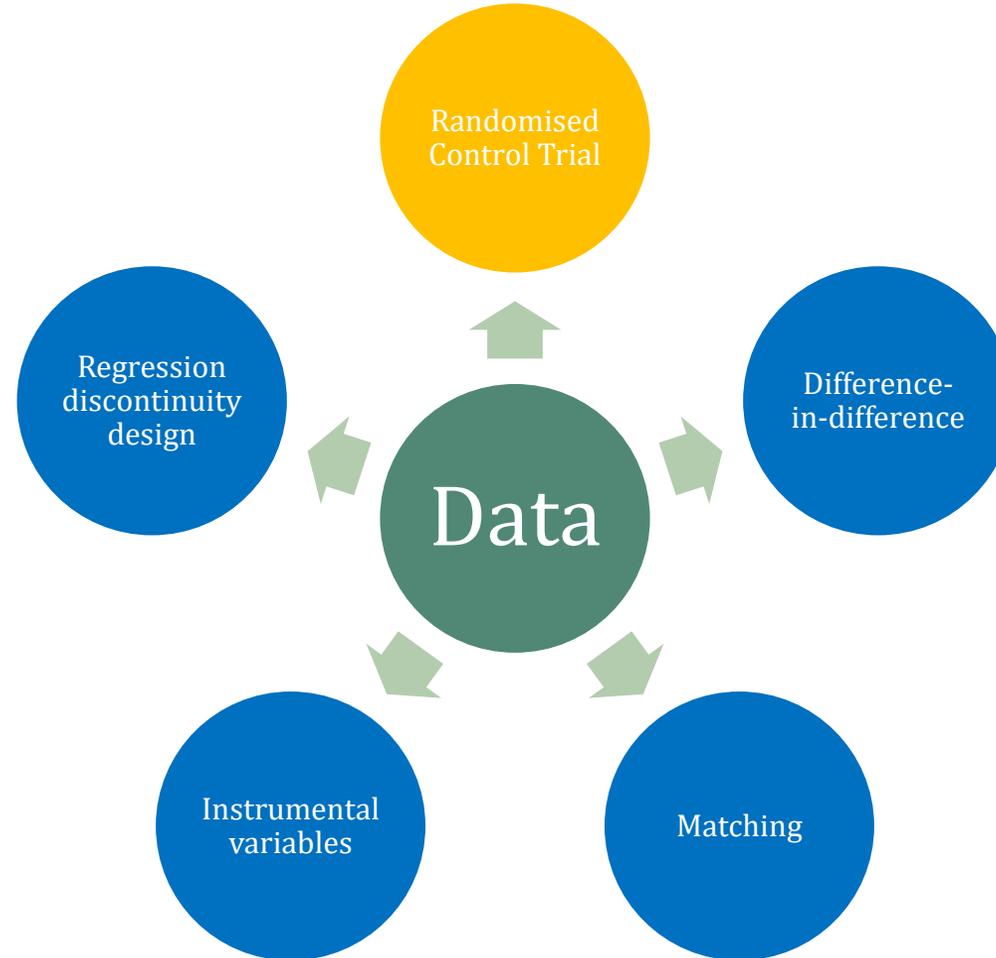
*“Even the best designed monitoring and evaluation systems will collapse without high quality data underpinning them.”*

Gugerty & Karlan (2018)

# Relevance of Data Collection in CIE



# Relevance of data collection in CIE



# Relevance of data collection in CIE

---

- To answer various evaluation questions and to measure impact we require data
  - Does the program improve the income of beneficiaries?
    - Requires data on the income of treatment and comparison groups
  - Does the program improve the resilience of beneficiaries?
    - Requires data on resilience of treatment and comparison groups
- Well chosen and well implemented methods for data collection and analysis are essential

# Relevance of data collection in CIE

- Poor quality data can undermine even the best designed CIE.

“Garbage in, garbage out”



Your analysis is as good as your data.

# Mixed Methods Approaches

---

- Mixed methods CIE is increasingly becoming common
- Understand the views, experiences and motivations of beneficiaries, implementers and stakeholders in greater depth.
- Understand the processes and mechanisms by which impacts occur – How and why?
- Investigate if a project had any unintended (both positive and negative) consequences
- Questions about meaning and motivation examine how a particular behavior or action is understood, or how people make sense of their circumstances.

# Data Collection for Mixed Methods Approaches

---

- To conduct a mixed methods approach the evaluator has to collect both:
- Quantitative Data
  - Ultimately numbers and statistics
  - “What is your income this month”
  - “On a scale of 1-10 how satisfied are you with your life
- Qualitative Data
  - Words, pictures, objects: Interpretation of Data
  - “Take me through a typical day when looking for work”

# Quantitative and Qualitative Data

## Quantitative Methods



## Qualitative Methods



# Qualitative Data Sources

---



**Key Informant Interviews** with people (e.g. community leaders, program staff) who have particularly informed perspectives on an aspect of the program being evaluated

**In-depth interviews** with participants to learn about about their experiences and expectations related to the program,



**Focus group discussions** are group interview designed to explore people's attitudes about aspects being evaluated



**Observations:** systematic observations to understand phenomena , especially hidden ones (e.g. child labour)

# Relevance of data collection in CIE

---

- The relevance of data collection in a CIE is not limited to indicators of impact:
  - Strong monitoring data can help to assess various parts of the Theory of Change and test assumptions.
  - If effectively integrated into the CIE, monitoring data can even be included in the model for estimating impact.
  - Relevant secondary data can be used to benchmark results against.
  - Data on characteristics that may influence program participation or outcomes (control variables)
    - Can be useful for various reasons depending on the approach
    - E.g. Capturing the education level of study participants for a employment programme Assists in assessing the robustness of a counterfactual

# Breakout Session

---



As a group - outline a basic causal chain for the EUTF program you work on:

- Otherwise you can select an intervention type
  - Vocational Training for Unemployed Youth
  - Access to Social Loans for Refugees
- Go From Input to Outcome
  1. For the causal chain consider three relevant evaluation questions or questions that could be asked to check if the assumptions in the chain hold.

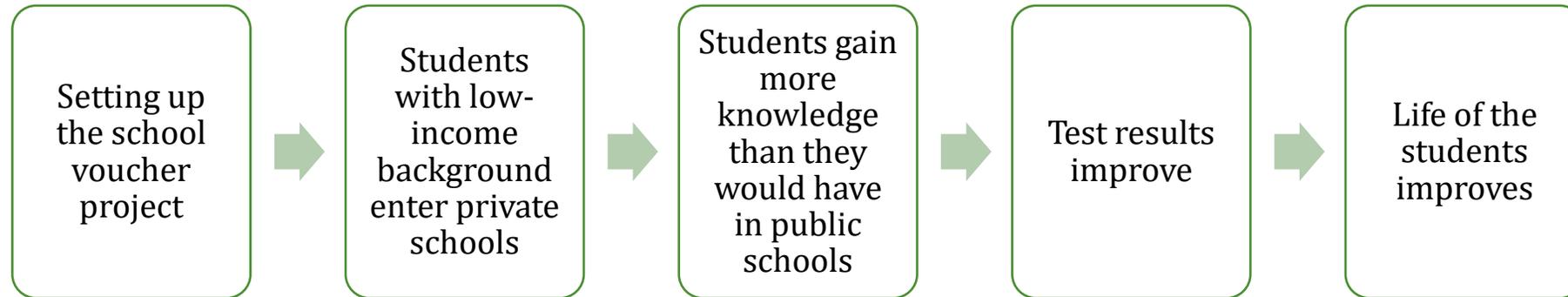
# Breakout Session

---

Consider the potential data that could be used to answer the question

1. Is the question best answered using quantitative or qualitative data, or both?
2. Consider the potential source(s) of the data. Consider:
  - a. Does the data already exist in some form? (Secondary Data)
  - b. Does the data need to be collected specifically for the evaluation?  
(Primary Data)
3. Can you see any challenges or limitations to using the data to answer the questions? (consider anything- Practical, technical)
4. Each team will present their table and go through one example in it (1 minute per team) at the end for the rest of the group

# Example – School Voucher

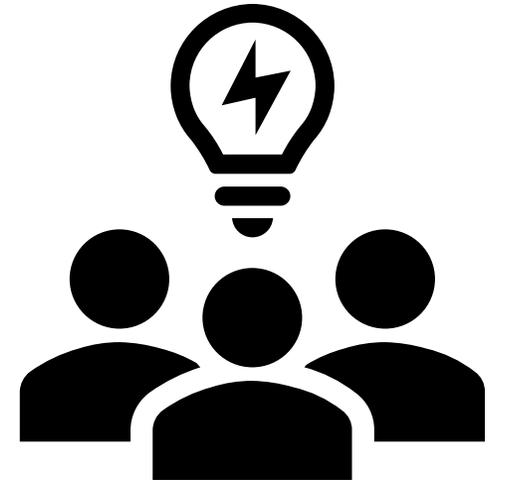


	Question (Evaluation or Testing Assumption)	Brief Description of Data that can be used to answer the question	Type of Data (Qual / Quant)	Data Source(s)	Challenges
1	Were vouchers effectively targeted?	Using monitoring data to review how many vouchers were delivered and whether target was met	Quant	Monitoring data/Secondary	
		Experiences of implementing partners on how identifying appropriate students and delivering the vouchers	Qual	Key informant interviews/Primary	
2	Do vouchers lead to students with low-income backgrounds attending private school?	Registration and Attendance data in study schools linked with use of voucher	Quant	Administrative Data/Secondary	Relies on schools having accessible and reliable attendance data
		Discussing with parents' the desirability and feasibility of sending their child to a private school even with a voucher	Qual	Focus Group Discussion/Primary	
3	Do test results improve?	Conducting learning assessments with students to measure learning outcomes	Quant	Student Tests/Primary	Tests need to be appropriate to the curriculum
		-	-	-	

# Breakout Sessions

---

- We will now briefly send you to your breakout rooms.
- Please write a message in the chat if:
  - You have not been added to a breakout room (you will be assigned one)
  - You believe you have been added to the wrong breakout room
- Please nominate a group speaker from each team who will help to facilitate the tasks and report back.



# Breakout Sessions

---

- For the task you should complete the template that will be shared in the breakout room
- The speaker for the group should open the template, share their screen with the other participants in the room and write the input
- You will have 15 minutes for the task



**END OF SESSION 2**

# **Session 3a: Stages of Preparation for Data Collection in CIE**

## **Moving from Evaluation Questions to Indicators**

## Session 3a Learning Objectives

---

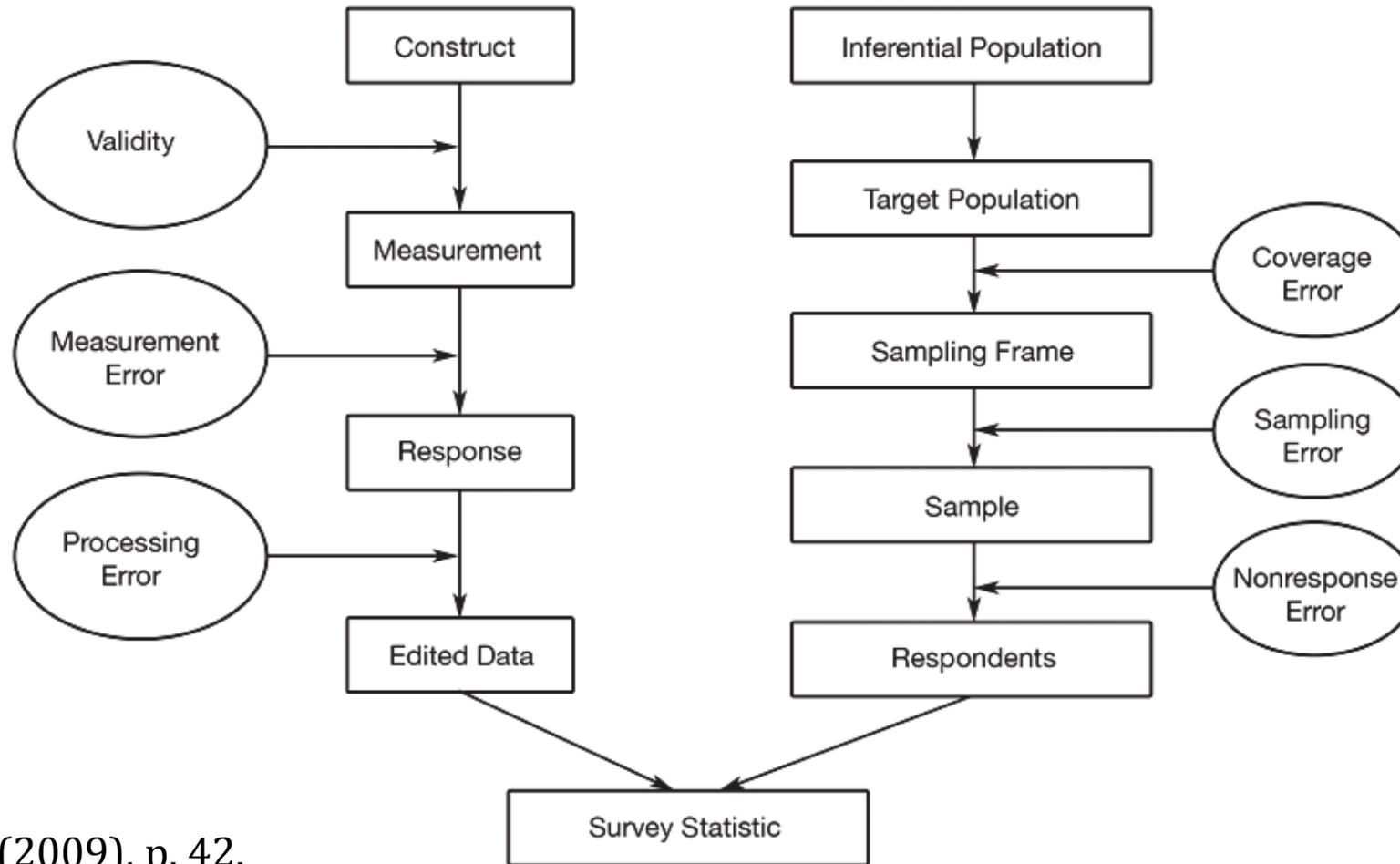
- Participants will be able to describe the steps in moving from Evaluation Question to collecting data to answer it
- Participants will be able to describe what is meant by a construct and an indicator when trying to answer an evaluation question
- Participants will be able to explain the importance of having clearly defined concepts when developing indicators
- Participants will be able to explain the process of using existing tools to develop questions that will provide relevant indicators

# Data Collection in CIE

---

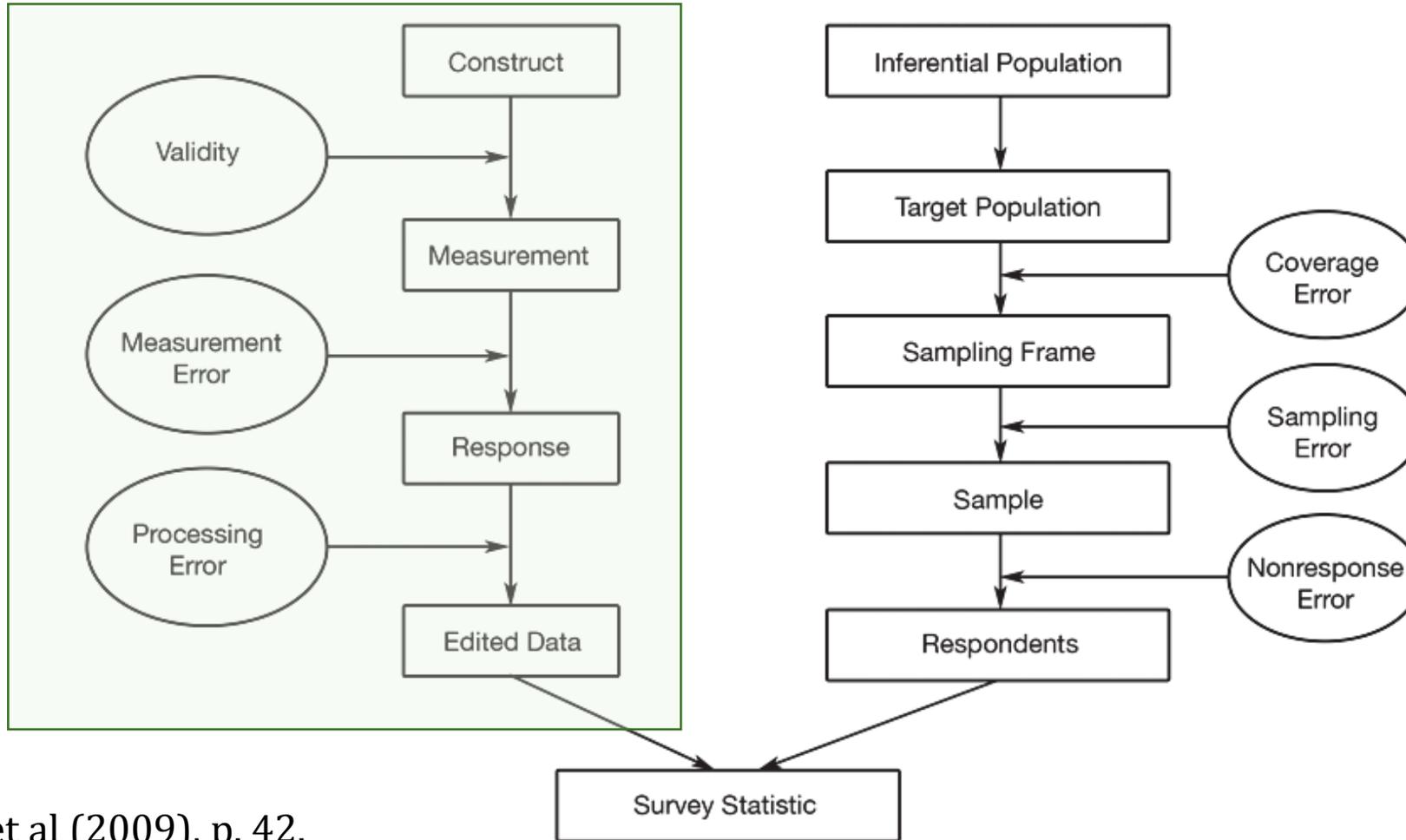
- In this workshop we will focus on collecting micro-data in hard-to-reach areas
- Hard-to-reach areas can mean various things:
  - Geographic - remote
  - Insecurity and fragile contexts
  - “Invisible” Populations
- Collecting data in hard-to-reach areas pose challenges and threats to CIE
  - Developing Tools
  - Sampling
  - Conducting Fieldwork

# Survey Life from a Quality Perspective



Groves et al (2009), p. 42.

# Survey Life from a Quality Perspective



Groves et al (2009), p. 42.

# Preparing for data collection

---

*.....To be well prepared when designing data collection, know your research questions and objectives.....*

# Evaluation Question to Data Collection



Defining concepts within  
the evaluation question

How can we try and  
measure this in the real  
world?

Questions within a survey

# Evaluation Questions

---

- What effects do the interventions have on livelihood in terms of economic wellbeing of refugee and host communities?
- What is the impact of the interventions on integration and social cohesion of refugees and host communities?
- How do we go from these evaluation questions to individual questions in a survey?

# Concepts of Measurement

---



What effects do the interventions have on livelihood in terms of economic wellbeing of refugee and host communities?

# Concepts of Measurement

---

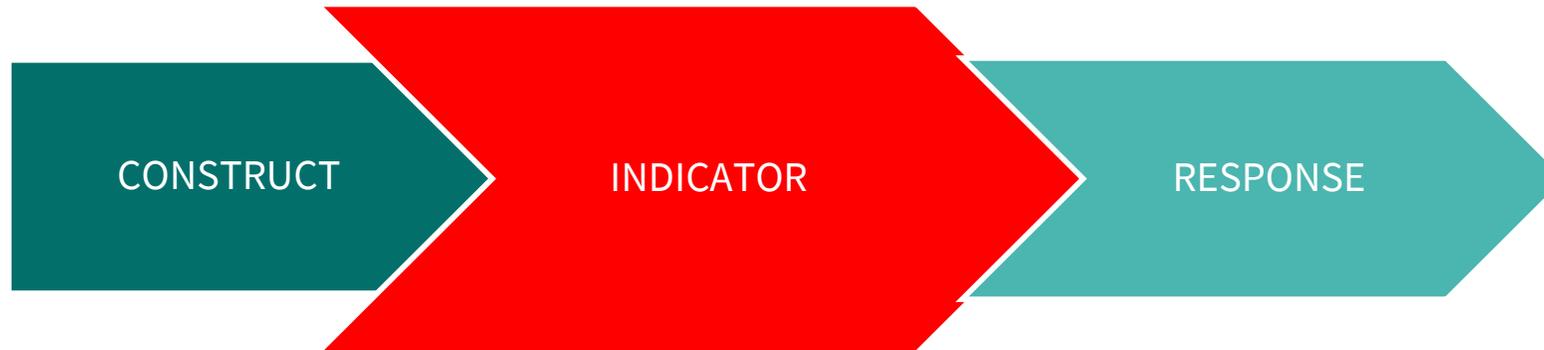


What effects do the interventions have on livelihood in terms of **economic wellbeing** of refugee and host communities?

- The ability of an individual to demand goods and services, in relation to its needs.

# Concepts of Measurement

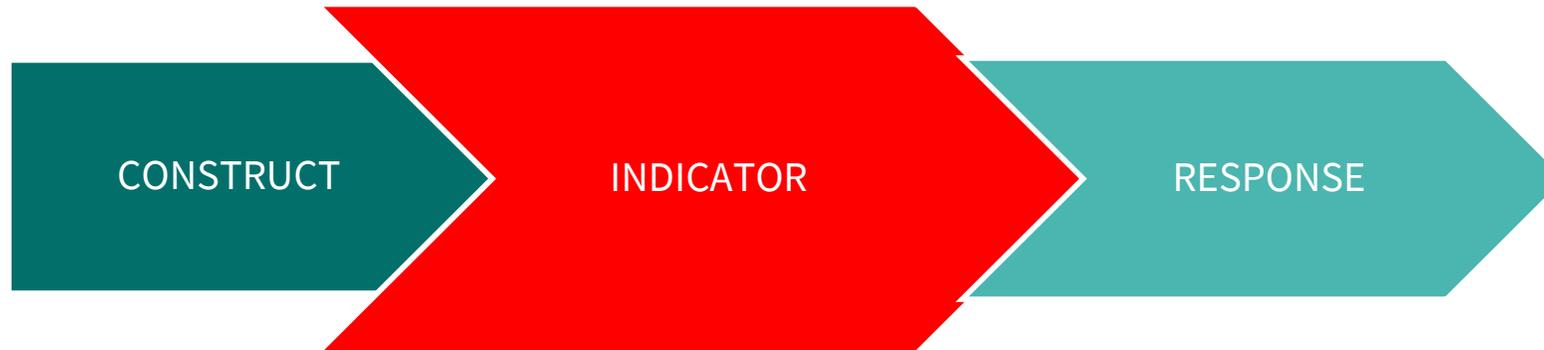
---



- Average income
- Employment Status
- Security in employment
- Business Ownership
- Business Performance
- Asset Ownership

# Concepts of Measurement

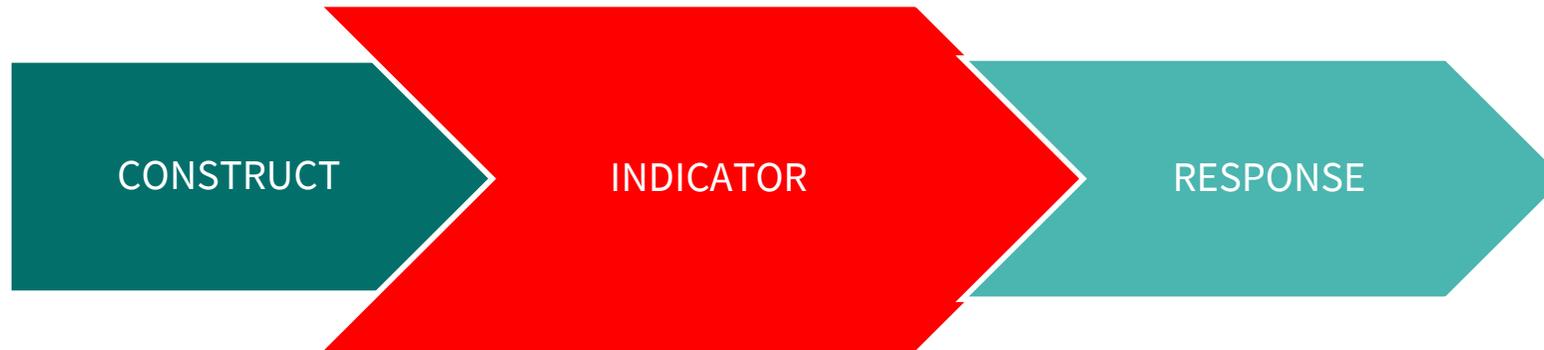
---



- Important to be clear and concise with your indicators as we you will always be limited with what you can collect
  - Which are the key indicators?
  - Which are most closely linked to the program?
  - Which are realistic to collect?

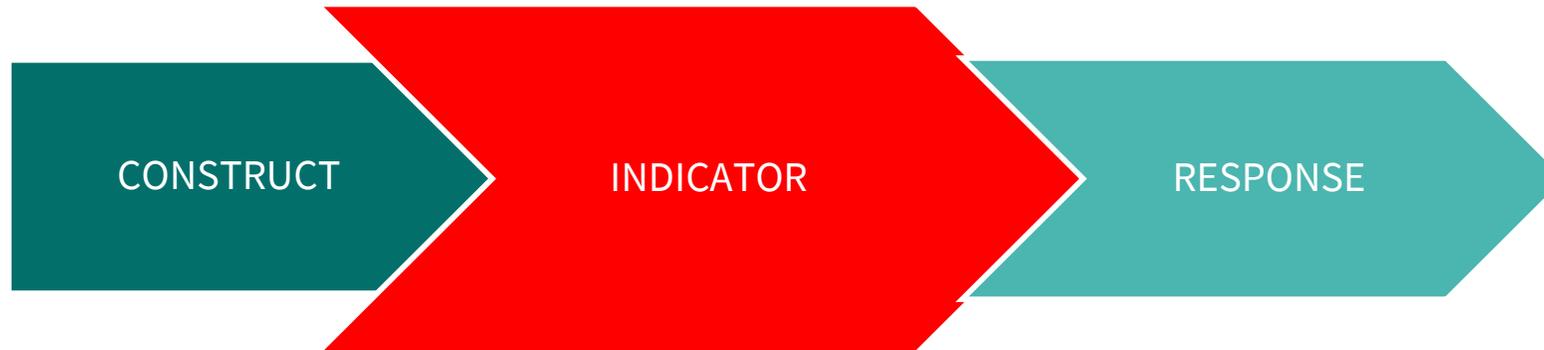
# Concepts of Measurement

---



- Average income
- Employment Status
- Security in employment
- Business Ownership
- Business Performance
- Asset Ownership

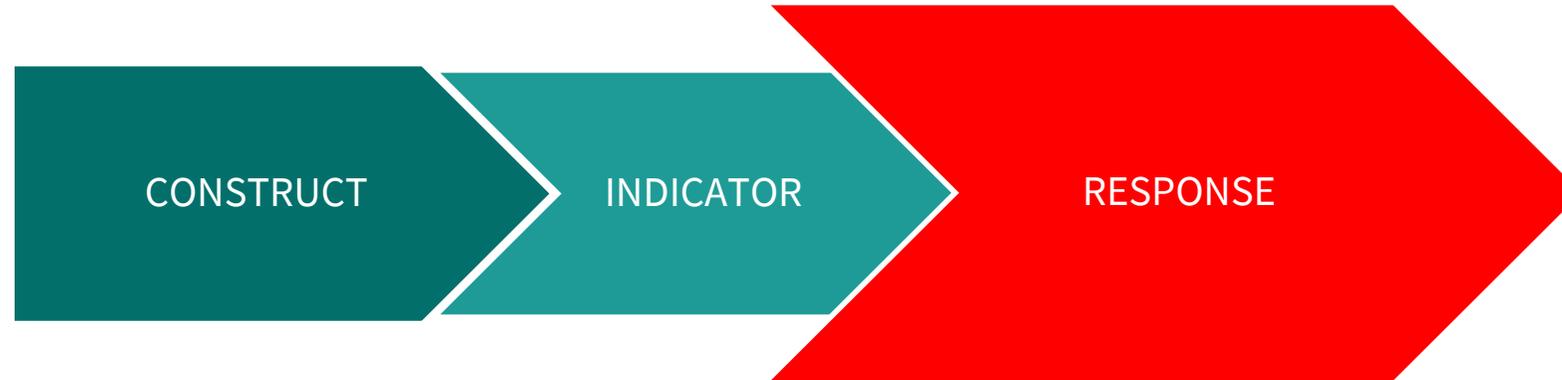
# Concepts of Measurement



- How exactly do we wish to define employment status?
- National / International definitions:
  - ILO definition of employed: employed person is a person aged 15 years or older who have worked (for pay or profit for at least one hour during a given week or having a job from which being absent under conditions on the reason of absence (holidays, sick leave, maternity leave, etc.) or duration

# Concepts of Measurement

---



sectionBsetup	Now I would like to ask you some questions about jobs or other income generating activity that you have or taken part in recently.																									
b1a (required)	b1a. In the last 7 days, did you do any work, even for just one hour? <i>This refers to any income generating job. Select all that apply</i>	<table border="1"> <tr><td>1</td><td>A paid employee of someone who is not a member of your household</td></tr> <tr><td>2</td><td>A paid worker on household farm or non-farm business enterprise</td></tr> <tr><td>3</td><td>An employer</td></tr> <tr><td>4</td><td>A worker non-agricultural own account worker, without employees</td></tr> <tr><td>5</td><td>Unpaid workers (e.g. Homemaker, working in non-farm family business)</td></tr> <tr><td>6</td><td>Unpaid farmers</td></tr> <tr><td>7</td><td>None of the above</td></tr> </table>	1	A paid employee of someone who is not a member of your household	2	A paid worker on household farm or non-farm business enterprise	3	An employer	4	A worker non-agricultural own account worker, without employees	5	Unpaid workers (e.g. Homemaker, working in non-farm family business)	6	Unpaid farmers	7	None of the above										
1	A paid employee of someone who is not a member of your household																									
2	A paid worker on household farm or non-farm business enterprise																									
3	An employer																									
4	A worker non-agricultural own account worker, without employees																									
5	Unpaid workers (e.g. Homemaker, working in non-farm family business)																									
6	Unpaid farmers																									
7	None of the above																									
b1b (required)	b1b. Do you have a <b>paid</b> permanent/long term job (eventhough you did not work in the last 7 days) from which you were temporarily absent? <i>Hint: This includes working for someone outside or inside the household, self-employed, military service and internships and apprenticeships, if paid.</i>	<table border="1"> <tr><td>0</td><td>No</td></tr> <tr><td>1</td><td>Yes</td></tr> </table>	0	No	1	Yes																				
0	No																									
1	Yes																									
b1c (required)	b1c. What is the main reason that you did not work in the last 7 days although you have a permanent job?	<table border="1"> <tr><td>1</td><td>Paid leave</td></tr> <tr><td>2</td><td>Unpaid leave</td></tr> <tr><td>3</td><td>Own illness</td></tr> <tr><td>4</td><td>Maternity leave</td></tr> <tr><td>5</td><td>Care of household member</td></tr> <tr><td>6</td><td>Holidays</td></tr> <tr><td>7</td><td>Strike/Suspension</td></tr> <tr><td>8</td><td>Temporary workload reduction</td></tr> <tr><td>9</td><td>Closure</td></tr> <tr><td>10</td><td>Bad weather</td></tr> <tr><td>11</td><td>School/Education/Training</td></tr> <tr><td>-96</td><td>Other (Specify)</td></tr> </table>	1	Paid leave	2	Unpaid leave	3	Own illness	4	Maternity leave	5	Care of household member	6	Holidays	7	Strike/Suspension	8	Temporary workload reduction	9	Closure	10	Bad weather	11	School/Education/Training	-96	Other (Specify)
1	Paid leave																									
2	Unpaid leave																									
3	Own illness																									
4	Maternity leave																									
5	Care of household member																									
6	Holidays																									
7	Strike/Suspension																									
8	Temporary workload reduction																									
9	Closure																									
10	Bad weather																									
11	School/Education/Training																									
-96	Other (Specify)																									
b1c_other	b1c_other. Please specify other																									

# Concepts of Measurement

---



What is the impact of the interventions on integration and social cohesion of refugees and host communities?

# Concepts of Measurement

---



What is the impact of the interventions on **integration** and **social cohesion** of refugees and host communities?

# Concepts of Measurement

---

- What is the impact of the interventions on **integration** and **social cohesion** of refugees and host communities?
  - Example of a hard-to-reach population
- What challenges do we face when going through the same process for these constructs



# Concepts of Measurement

- World Bank have developed a Social Capital and Social Cohesion Measurement toolkit

**Table 2: Social Cohesion - Key Dimensions and Definitions**

DIMENSION		DEFINITION	SOURCE
Social Cohesion		A sense of shared purpose and trust among members of a given group or locality and the willingness of those group members to engage and cooperate with each other to survive and prosper.	Stanley 2003; Chan et al. 2006; Mvukiyehe 2011 SIPA 2018
Trust		An individual's 1) belief that another individual, group, or institution that could do her harm or betray her will not do so and 2) willingness to take actions that make herself vulnerable to that actor.	Levi & Stoker 2000; Gambetta 2000; Ostrom & Ahn 2009; Gilligan, Pasquale, & Samii 2013; Scrivens & Smith 2013
Collective Action Norms		Collectively shared and internalized moral prescriptions that encourage costly actions that primarily benefit others.	Ostrom 1998; Ostrom 2005; Ostrom & Ahn 2009; Fehr & Fischbacher 2003; Benabou & Tirole 2005
Shared Purpose	Belonging	The degree to which an individual or collective group feel like they "fit" together in a group.	Pham & Vinck (UNICEF) 2017
	Identity	The characteristics that an individual or collective group believe to define them.	Pham & Vinck (UNICEF) 2017
	Attitudes Toward Out-Groups	How individuals perceive people with other values, lifestyles, or identities within their group or locality.	Larsen, Koch, & Dragolov 2013; Janmaat & Keating 2019; Bogardus 1925
Civic Engagement		The attitudes and behaviors of individuals that result in participation to improve local area conditions for others and/or help shape the area's future.	Adler & Goggin 2005

# Concepts of Measurement

---

- A set of 15 survey questions that measure the dimensions of social cohesion (Trust, Civic Engagement, Belonging etc.).
- However, prior to the quantitative data collection, qualitative data collection is used to understand the context and adapt the tool to be relevant to the evaluation and population.
- Toolkit contains:
  - Quantitative survey questions
  - Qualitative survey script

# Concepts of Measurement

---

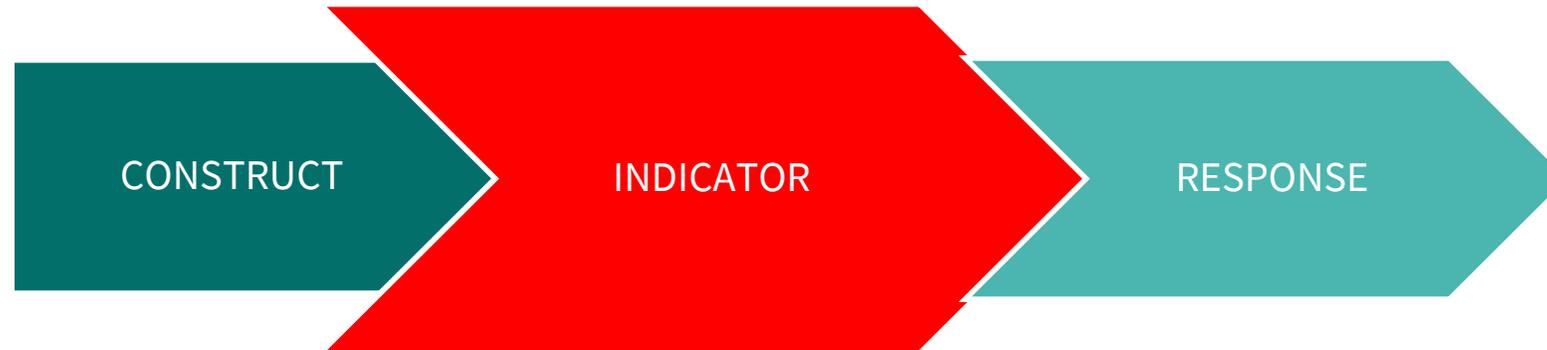


## Social Cohesion:

- Trust
  - An individual's 1) belief that another individual, group, or institution that could do her harm or betray her will not do so and 2) willingness to take actions that make herself vulnerable to that actor.

# Concepts of Measurement

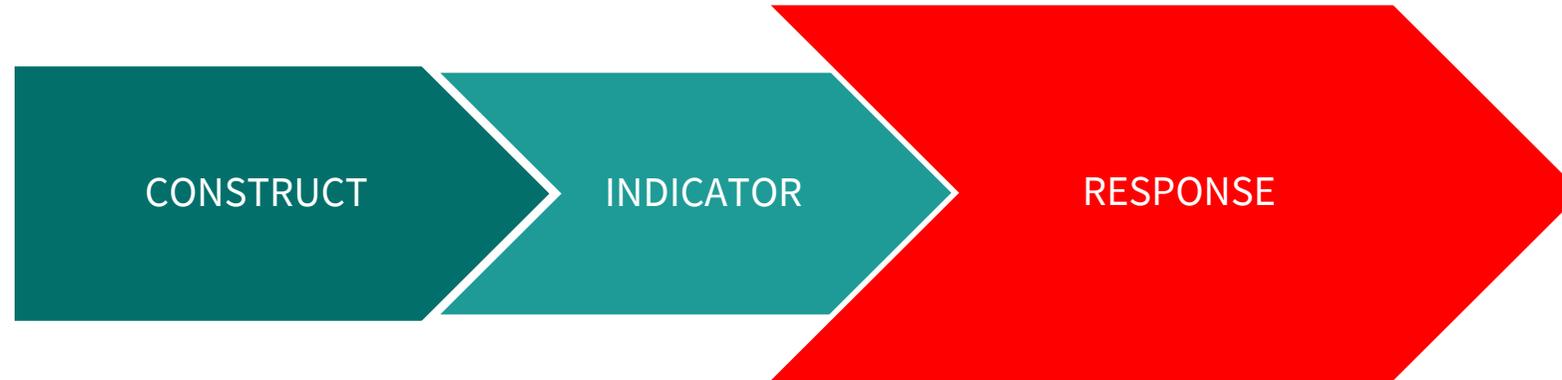
---



- Indicator for Trust:
  - Trust score using Social Capital and Social Cohesion Measurement toolkit

# Concepts of Measurement

---



Question	Response
<p>Please tell me the degree to which you agree or disagree with the following statement: The following types of people are likely to take advantage of you.</p> <p>a. Individuals from my [social group]</p> <p>b. Individuals from a different [social group]</p> <p>c. Individuals/organizations/ institutions representing linking relationships between social networks with differing levels of power or social status]</p>	<p>1. Strongly agree</p> <p>2. Agree</p> <p>3. Neither agree nor disagree</p> <p>4. Disagree</p> <p>5. Strongly disagree</p> <p>777. Do not know anyone from this group</p>
<p>Please tell me the degree to which you agree or disagree with the following statement:</p> <p>If I was at a [CDD geographic unit] meeting and accidentally left [my wallet] behind, I believe that the person who found it would return it to me.</p>	<p>1. Strongly agree</p> <p>2. Agree</p> <p>3. Neither agree nor disagree</p> <p>4. Disagree</p> <p>5. Strongly disagree</p> <p>777. Do not know anyone from this group</p>

# Developing Survey Tools

---

- Literature review of existing tools
- Example of repositories for surveys: World Bank, UNICEF, UNDP, JPAL
- For indicators of more abstract concepts the tools should have been tested before
- Consider the context – often well established indicators will have been tested in various contexts

# **Session 3b: Stages of Preparation for Data Collection in CIE**

## **Developing Data Collection Tools**

## Session 3b Learning Objectives

---

- Participants will be able identify good and bad practices when developing survey questions
- Participants will be able to understand the benefits and approaches to pre-testing survey tools
- Participants will be able to explain what measurement error is
- Participants will learn approaches to avoid measurement error in when collecting data in hard-to-reach areas

# Developing Survey Questions

---

- Best to not try and ‘re-invent the wheel’ and use existing tools when possible
- However not always possible:
  - Programme specific questions
  - Basic monitoring outcomes
  - Contextualising existing tools

# Developing Survey Tools

---

- **Guiding principles for question wording**
  - **Brief:**
    - Unless a word is relevant to the question, it should not be included
    - Long questions tend to be overly complex
  - **Be objective:**
    - Pay attention Neutrality of words used in the question
  - **Be simple:**
    - Use the simplest of words and phrases – avoid technical jargon and acronyms
  - **Be specific**
    - Ask very precise questions – use clearly defined concepts and reference periods

# Developing Survey Tools

---

- **Guiding principles for question wording**
  - **Brief:**
    - Unless a word is relevant to the question, it should not be included
    - Long questions tend to be overly complex
  - **Be objective:**
    - Pay attention Neutrality of words used in the question
  - **Be simple:**
    - Use the simplest of words and phrases – avoid technical jargon and acronyms
  - **Be specific**
    - Ask very precise questions – use clearly defined concepts and reference periods

# Be Objective

---

- **Loaded Question:** Question that makes an assumption on the behalf of the respondent
  - Example: *Has your improved self-confidence from taking part in this training helped you find a job?*
- **Leading Questions:** Questions that push respondent in the direction of a certain answer
  - Example: *Don't you think your employability has increased since you completed this training?*

# Developing Survey Tools

---

- **Guiding principles for question wording**
  - **Brief:**
    - Unless a word is relevant to the question, it should not be included
    - Long questions tend to be overly complex
  - **Be objective:**
    - Pay attention Neutrality of words used in the question
  - **Be simple:**
    - Use the simplest of words and phrases – avoid technical jargon and acronyms
  - **Be specific**
    - Ask very precise questions – use clearly defined concepts and reference periods

# Poor Practice in Question Writing

---

- Indefinite words
  - Example: *What is your educational background?*
- Abbreviations
  - Example: *Which VTI Did you attend for the EUTF training?*
- Double barrelled Questions
  - Example: *Did you attend and complete the skills training on tailoring?  
(Yes/No)*

# Developing Survey Tools

---

- **Guiding principles for question wording**
  - **Brief:**
    - Unless a word is relevant to the question, it should not be included
    - Long questions tend to be overly complex
  - **Be objective:**
    - Pay attention Neutrality of words used in the question
  - **Be simple:**
    - Use the simplest of words and phrases – avoid technical jargon and acronyms
  - **Be specific**
    - Ask very precise questions – use clearly defined words and concepts and reference periods

# Poor Practice in Question Writing

---

- Indefinite words
  - Example: *What is your educational background?*
- Reference period:
  - Example: *Have you taken out a loan recently?*

# Poor Practice in Question Writing

---

- In your hair salon business, what was your total profit in 2020?
  - Long recall period given required information
- How much did your employers spend on market research in the last 12 months?
  - Not relevant to the respondent and unlikely to provide information
- Have you ever been a victim of human trafficking?
  - Potentially very sensitive information



# Questionnaire Development

---

- Alongside the writing of individual questions, care should be taken when considering the structure of the questionnaire
- Questionnaire Flow:
  - Logical flow of ideas and reasoning of respondents
    - E.g. Take a list of jobs in the last 6 months and one-by-one go through them with follow up questions
    - At the start of the section explain the theme and structure of the questions
  - Put the respondent at ease
    - Avoid putting sensitive or cognitively challenging questions at the start of the questionnaire
- Questionnaire Length:
  - Too long and can lead to fatigue and affect data quality
  - No strict timeline but
  - Need to be clear upfront with the length of the questionnaire

# Developing Survey Tools

---



*Drafting a perfectly worded survey while sitting in the office is near impossible.*

# Pre-testing

---

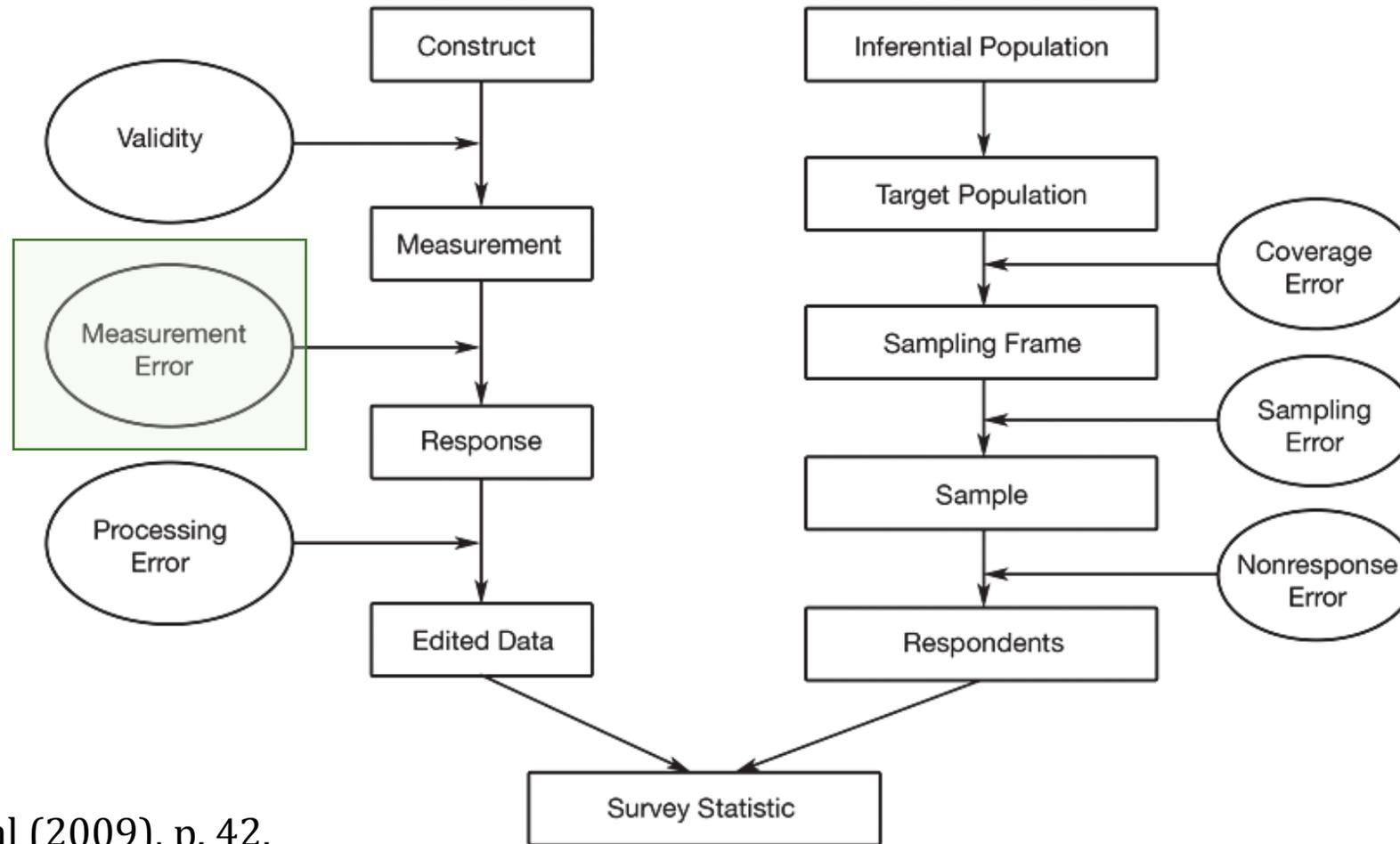
- It is important to pre-test any survey prior to launching
  - Whether it is an existing tool or newly written question it is important to test them with the target population
- Questions that respondents are struggling to understand or answer
  - Probe – why is this the case?
- Additional answer options
  - Are respondents giving answers that are not falling into the available categories?
- Questionnaire Length
  - Note: a good rule of thumb is that length will reduce from testing to the end of fieldwork by roughly 1/3<sup>rd</sup>
- Benchmarking data collected against expectations

# Pre-testing

---

- No set way of conducting a pre-test
- However it should contain a real life test of the survey with respondents similar to those targeted
- Often useful to workshop the tool prior to the field test
  - Knowledgeable participants on the theme and context to discuss the tool
  - People with a similar profile to the target respondents
  - Those administering the pre-test
- If budget allows for a large number of pre-test interviews, the researcher can also check psychometric properties
  - Internal consistency and validity

# Survey Life from a Quality Perspective



Groves et al (2009), p. 42.

# Measurement Error

---

- Measurement error is the difference between the value collected during a data collection and the ‘true’ value.
- Everything that we have discussed previously are potential threats to measurement error
  - Poorly designed questions and survey
  - Cognitive challenges in answering the question
  - Social desirability bias

# Measurement Error in Practice

---

- We will now take a bit of time to understand a couple of common challenges in measurement when conducting CIEs
- We will review why they occur and potential steps that can be taken to solve them

# Measurement Error in Practice

---

- In remote rural Africa majority participate in agriculture in smallholder farms
- Labor of household members in agriculture is a key asset for poor households
- Output of the labor is important in understanding the income of these households.
- However, they are particularly vulnerable to measurement error

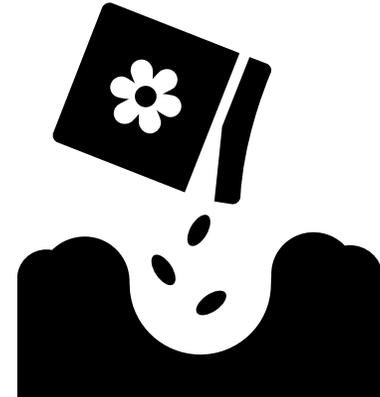
# Measurement Error in Practice

- Study of labor input during the rainy season in Mara Region of rural northern Tanzania (Arthi et al., 2019)
- Households were randomly assigned to one of four survey designs:
  - Weekly face-to-face survey
  - Weekly phone survey
  - Face-to-face survey at the end of rainy season
  - Face-to-face survey at the end of rainy season where labor time broken down by activity



# Measurement Error in Practice

- In the respondents that recalled their labor at the end of the season, the total number of hours worked was three times higher than those with weekly surveys.
- Recall failure



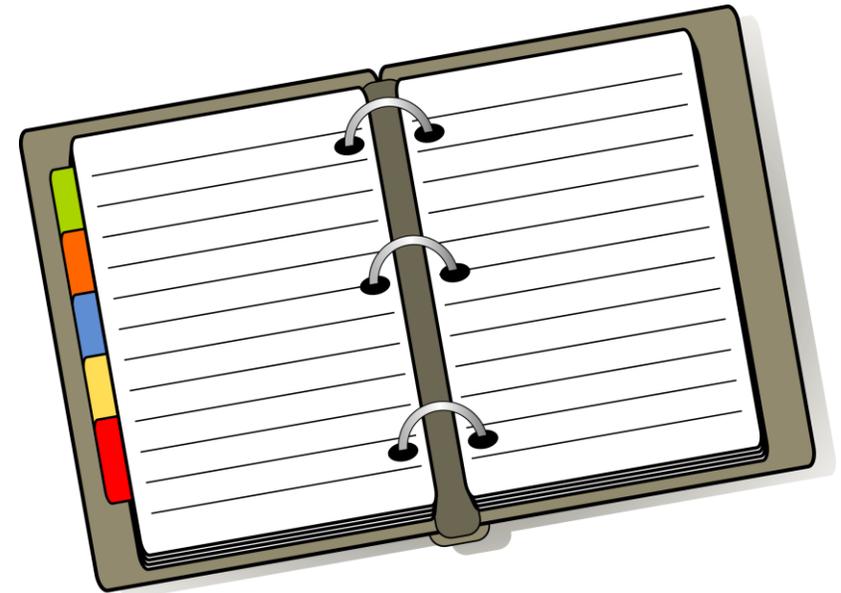
# Measurement Error in Practice

- Asking farmers to self-reporting agricultural production often requires a lot of guesswork and often leads to underestimation (Desiere and Jolliffe, 2018).
- Even among high-quality surveys agricultural production estimates vary a lot



# Measurement Error in Practice

- One solution may be in asking respondents to keep a production diary simply asking them to self-report their production.
- This was tested alongside a self-reported survey of agricultural production in Uganda (Deininger et al. 2012).
- Farmers were provided with diaries and were visited by locally respected monitors



# Measurement Error in Practice

---

- It was found that there were substantial differences between the self-reported production and the production captured by the diary.
- Diary estimates of production are consistently higher and more in-line with expectation based on what we know about home consumption plus sales in households.
- No apparent issue in lower educated farmers completing the diary with supervision.

# Measurement Error in Practice

---

- Another common source of measurement errors is when respondents are asked to provide information on sensitive subjects.
  - Sexual behaviour
  - Income Information
  - Criminal Involvement
- Vulnerable to refusals or providing untruthful information
- Enumerators may also feel embarrassed to ask
- Technology can be utilised to minimize measurement error (more on this tomorrow)

**END OF SESSION 3**

# **Session 4a: Sampling in data collection: Challenges and implications.**

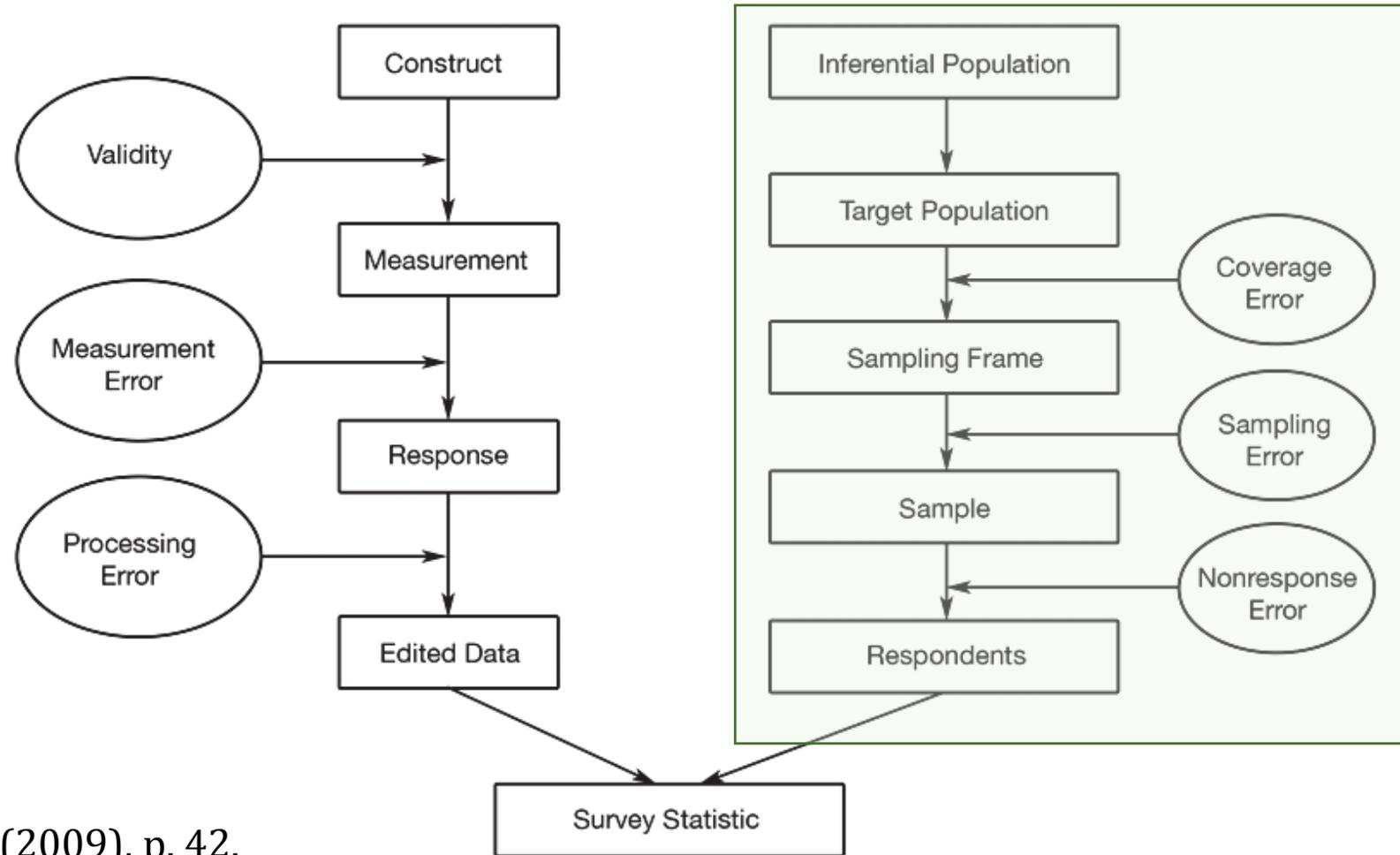
## **Developing a Sample Frame**

## Session 4a Learning Objectives

---

- Participants will be able to describe what a sampling frame is and it's relevant to a CIE
- Participants will be able to describe different strategies to use, in hard-to-reach-areas that may not have sampling frames, develop sample frames
- Participants will be able to explain the concept of **coverage error** and it's implications for research
- Participants will be able identify examples of coverage error

# Survey Life from a Quality Perspective



Groves et al (2009), p. 42.

# Definitions

---

- **Target population** = the group of elements for which the survey makes inferences
- **Sampling frame** = lists or procedures used to identify all elements of the target population
- **Sample** = the group of elements of the frame from which measurement will be sought
- **Respondents** = elements that are successfully measured

# Sampling Frame

---

- The sample frame ideally includes all units from the population that the evaluation is focused on.
- It is from the sample frame that the sample will be drawn
- In CIE, this is usually a list of all the units that:
  - Received the program
  - Did not receive the program and are identified as the counterfactual group



# Sampling Frame

---

- Simple lists of population elements
  - Registries of persons (e.g. civil registry, voter registry)
  - Population census
  - Members of professional organizations
  - Business establishments registered in a particular city
  - Telephone book
- Maps of areas in which elements can be found (provinces, districts, villages)
- GIS data: location of buildings, fields, factories, etc.
- Time periods during which target events would occur (satisfaction surveys after zoo visit)

# Developing a Sampling Frame

---

- In hard-to-reach areas it may be that no sampling frame exists.
- Consider, for example, that you wish to understand the impact of improved mobile money on household finances in rural areas.
- You may have identified the villages included in the study using implementation data and administrative data.
- After calculating the required sample size, you need to interview 10 households in the village. But there is no list of all households of these village.
- Survey staff must develop a list of all HH which serves then as sampling frame

# Developing a Sampling Frame

- Practically there are different approaches that can be taken to obtaining a household list for a village.
- **Door-to-door listing:** Enumerators move door-to-door to every household in the village with the help of a knowledgeable village guide and list each household



# Developing a Sampling Frame

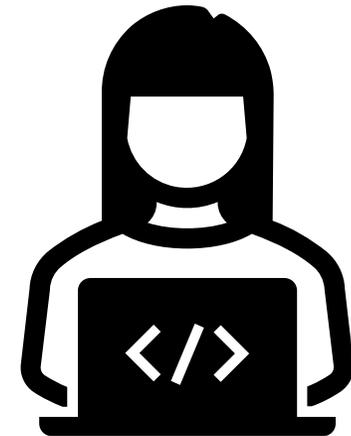
- When administrative data is not possible to use as a sampling frame it may be possible to use geospatial mapping.
- Often the case with hard-to-reach areas such as refugee settlements where it can be challenging to practically define divisions.
- Google Maps and mapping data can be used to identify groupings of households and act as a sample frame



Source: UNHCR.

# Developing a Sampling Frame

- The previous examples are when a sampling frame is developed well before the sampling takes place (i.e. which households are selected).
- This is the most rigorous approach and sometimes may be necessary depending on the sampling strategy such as if stratification is required (discussed later).
- **The sampling in this case can be conducted by the evaluator using statistical software and is transparent.**



# Developing a Sampling Frame

---

- For hard-to-reach areas therefore it can either be difficult and costly to acquire a sampling frame beforehand.
- What can we do if there are not the resources to do so?
- **Approach 1:**
  - 10 interviews are required in each village
  - Each enumerator can do two interviews per day.
  - A team of five enumerators are told to visit households one-by-one until they've completed two interviews each.
- What problems can you see with this approach?

# Developing a Sampling Frame

---

- This approach is not random and could be open to bias:
  - If all enumerators start village center then the households interviewed are likely to be in the vicinity of the village office
  - Enumerators may avoid travelling to harder to reach households
  - Enumerators may focus on more welcoming members of the community
  - Even if you tell the enumerators to go to houses “at random” – what does this mean?

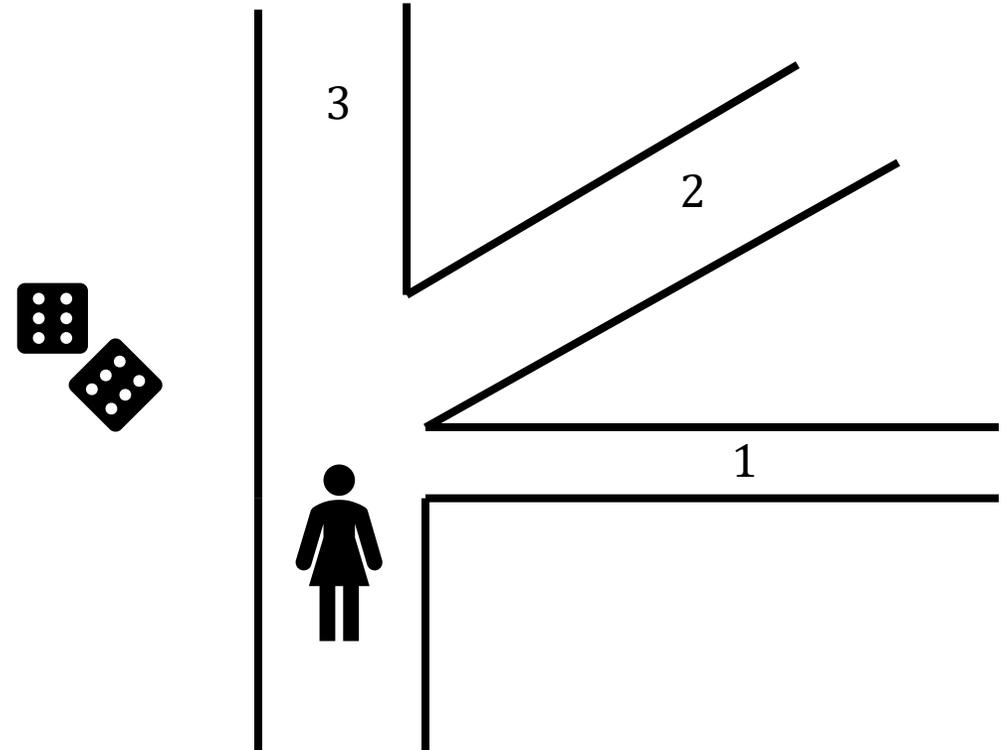
# When No Sampling Frame is Available

---

- You can aim to minimise this by having a process that is designed to randomly sample.
- This is often referred to as a random-walk approach

# When No Sampling Frame is Available

- Step 1: All enumerators go to the village center and each spins a bottle to determine the direction of travel.
- Step 2: At any intersection the enumerator assigns each direction with a dice
- Step 3: The dice is rolled until one of the directions is rolled.
- Step 4: Each random walk will last for 5 minutes (adjusted depending on size of village and number of households selected) and the next house after the time is selected
- **The enumerator cannot turn around until getting to the boundary or completing their interviews**

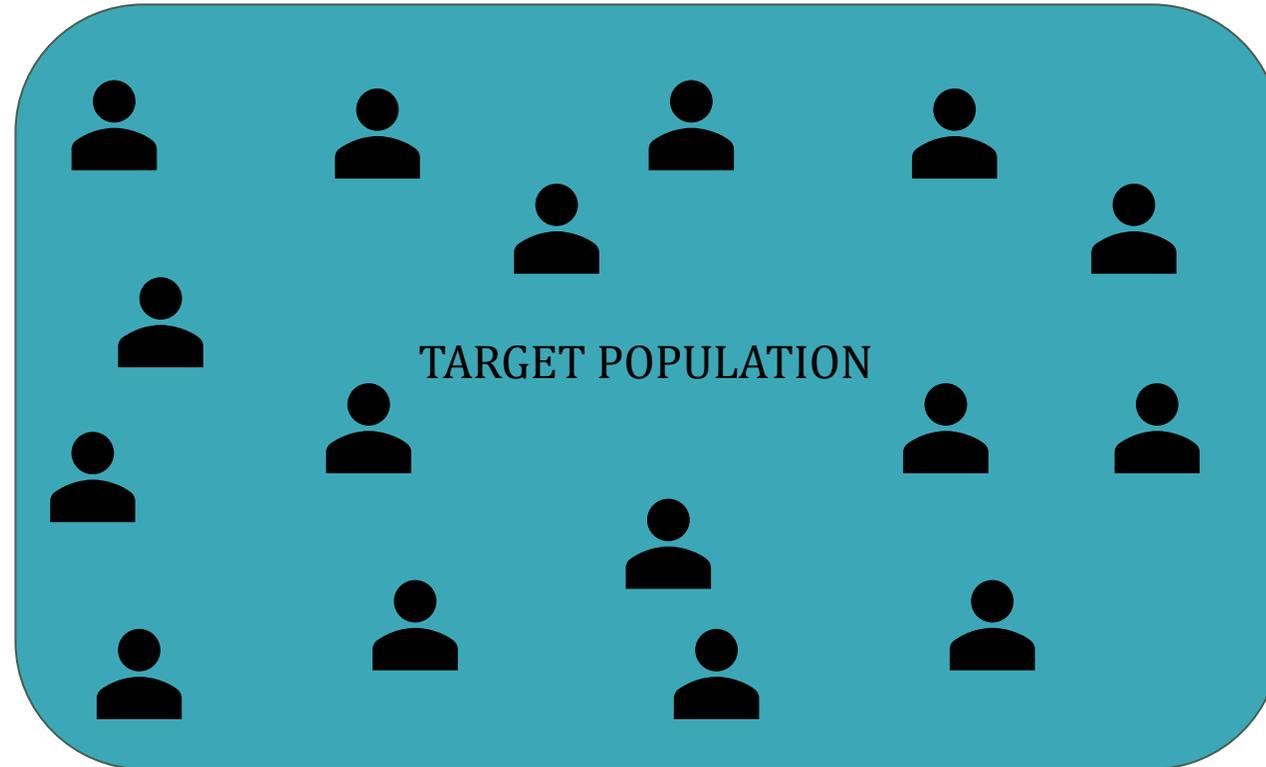


# Coverage Error

---

- Coverage error is an error that occurs when moving from identifying the target population to obtaining a sampling frame.
- An obvious example is using a city telephone registry as a sampling frame to understand public opinion of registered voters on an issue.

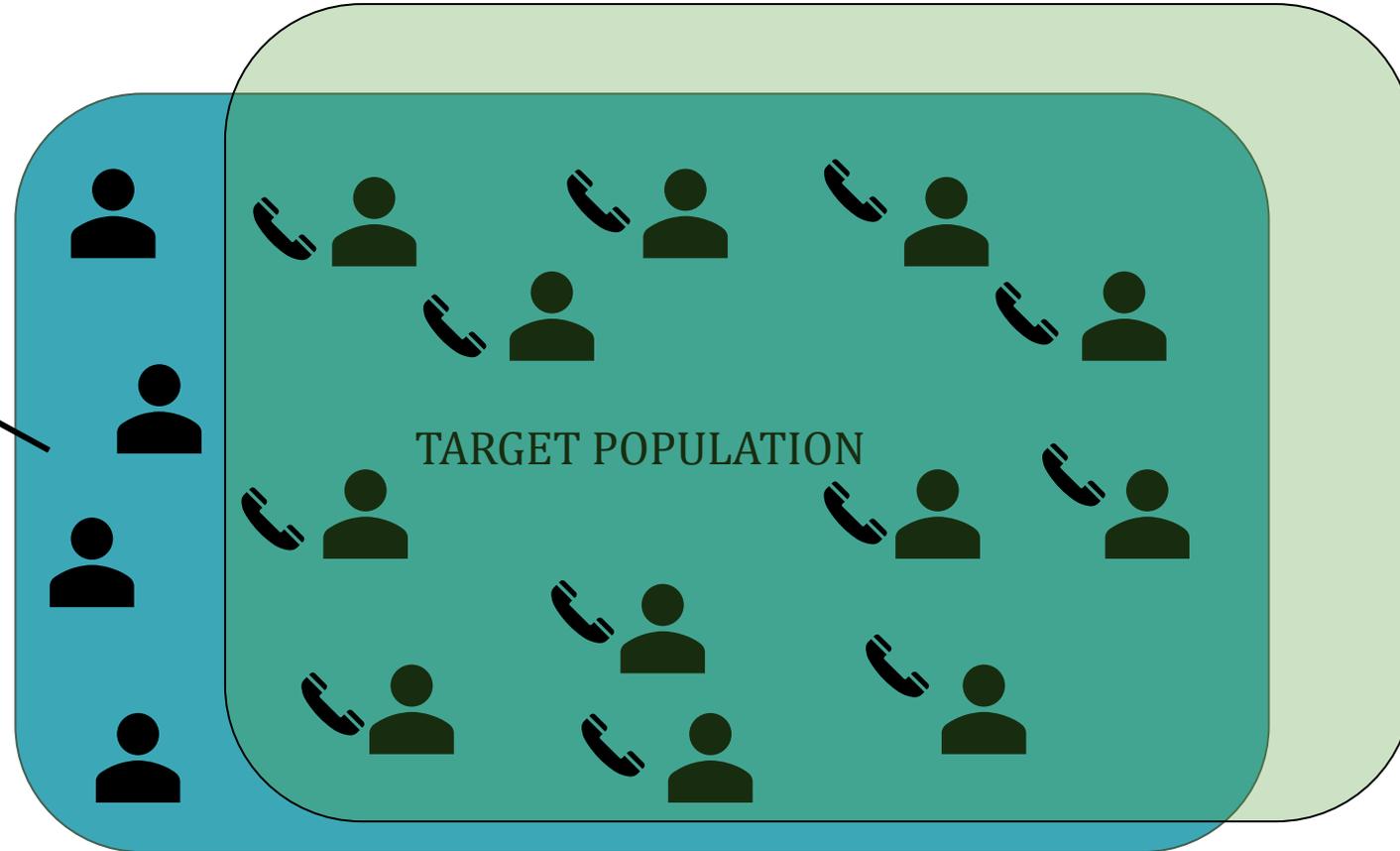
# Coverage Error



Example: households of a certain city

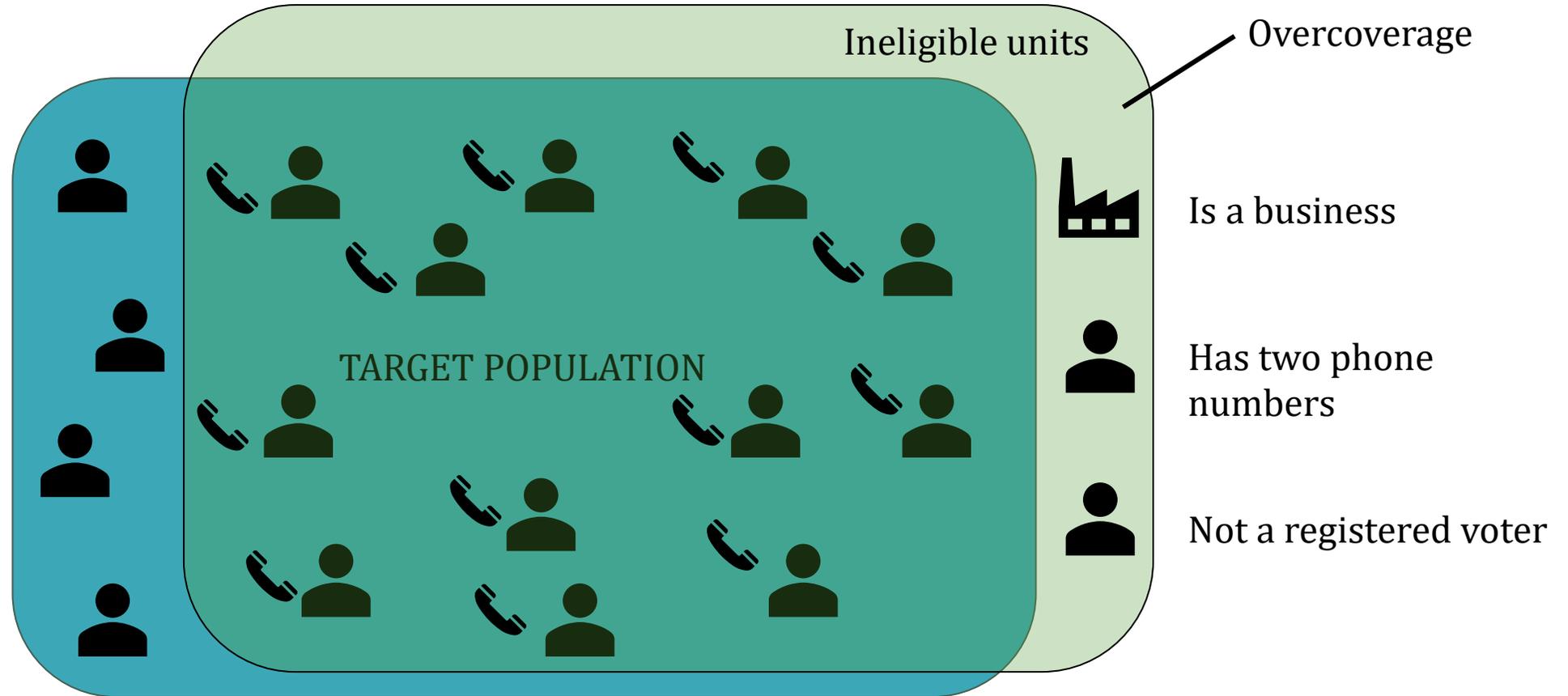
# Coverage Error

Undercoverage:  
People without  
landlines



Sampling Frame: telephone register of that city

# Coverage Error



Sampling Frame: telephone register of that city

# Coverage Error

---

- **Undercoverage:** some elements of the target population do not appear in the sampling frame
- **Ineligible units:** some units of the sampling frame are not part of the target population
- **Duplication:** several frame units are mapped onto one single element in the target population
- **Clustering:** multiple elements of the target population are linked to the same frame element

# Coverage Error

---

- If we return to the example where an evaluator wants to conduct household interviews in a village.
- They are provided with a list of households by the village head and that is used as the sampling frame for that village.

# Coverage Error

---

- Aside from the potential issues in estimating impact from coverage error it can also cause a lot of issues during fieldwork
- These problems during fieldwork can then filter down to cause other types of errors that affect impact estimates
  - Ineligible people being interviewed and included in analysis
  - Duplicate entries or non-existent entries being sampled and leading to a reduction in sample size if list not validated

# Coverage Error

---

- Important to have a sampling frame that is complete, valid and reliable
- Often for program CIE the sampling frames heavily utilise program data
  - Application database and selection information
  - Information on replacements or drop-outs of the program
  - Contact information for participants
- To ensure that an evaluator can use these effectively to build a sample frame
  - Integrated into a centralised monitoring system to bring together documents used in field
  - Easy to link various documents (i.e. Unique ID for each participant)
  - As up-to-date as possible
  - Basic quality assurance checks – no duplicates, totals match

# **Session 4b: Sampling in data collection: Challenges and implications.**

## **Taking a Sample**

## Session 4b Learning Objectives

---

- Participants will be able to describe using randomization to take an unbiased sample
- Participants will be able to understand how the context, needs and implementation of their program affects the required sample size for a CIE
- Participants will be able to describe the necessity for sufficient sample size in the credibility of a CIE

# Sampling

---

- After obtaining a sampling frame, it is often not feasible or efficient to collect data from everyone.
  - Therefore, a sample is drawn from the whole sampling frame.

**How should you draw a sample?**

# Sampling

---

- It depends on what questions you are trying to answer with your data.
- If you are trying answer questions about the population or program quantitatively:
  - What is the typical level of education in the participants of my program?
  - What is the rate of employment of participants after graduation
  - What is the impact of the program on income?
- You want to aim to draw an unbiased sample to get the best estimate of the population
- Generally, the bigger the sample size, the better the estimate

# Random Sampling

---

- Random sampling methods reduce the possibility of bias as the possibility of someone being selected as part of the sample relies completely on chance.
- Simple random sampling even can be done using Excel and using the =RAND() formula

ID	name	random number
1	Participant 1	0.380302351
2	Participant 2	0.30281822
3	Participant 3	0.887916562
4	Participant 4	0.114299885
5	Participant 5	0.222928459
6	Participant 6	0.1396707
7	Participant 7	0.981601282
8	Participant 8	0.45712799
9	Participant 9	0.219362055
10	Participant 10	0.195480155
11	Participant 11	0.729836065
12	Participant 12	0.570333368
13	Participant 13	0.039241158
14	Participant 14	0.3224227
15	Participant 15	0.902078485
16	Participant 16	0.483841749
17	Participant 17	0.779256552
18	Participant 18	0.696729917
19	Participant 19	0.660365209
20	Participant 20	0.731995707
21	Participant 21	0.70413838
22	Participant 22	0.879124571
23	Participant 23	0.633449236
24	Participant 24	0.84291115
25	Participant 25	0.757965883
26	Participant 26	0.951747111
27	Participant 27	0.300895972
28	Participant 28	0.016725954
29	Participant 29	0.249501224
30	Participant 30	0.193940095



ID	name	random number
28	Participant 28	0.016725954
13	Participant 13	0.039241158
4	Participant 4	0.114299885
6	Participant 6	0.1396707
30	Participant 30	0.193940095
10	Participant 10	0.195480155
9	Participant 9	0.219362055
5	Participant 5	0.222928459
29	Participant 29	0.249501224
27	Participant 27	0.300895972
2	Participant 2	0.30281822
14	Participant 14	0.3224227
1	Participant 1	0.380302351
8	Participant 8	0.45712799
16	Participant 16	0.483841749
12	Participant 12	0.570333368
23	Participant 23	0.633449236
19	Participant 19	0.660365209
18	Participant 18	0.696729917
21	Participant 21	0.70413838
11	Participant 11	0.729836065
20	Participant 20	0.731995707
25	Participant 25	0.757965883
17	Participant 17	0.779256552
24	Participant 24	0.84291115
22	Participant 22	0.879124571
3	Participant 3	0.887916562
15	Participant 15	0.902078485
26	Participant 26	0.951747111
7	Participant 7	0.981601282

# Sampling

---

- If you trying answer questions using qualitative data such as exploratory questions or understanding why something happened at a deeper level such as:
  - Why did participants not attend the program after being selected?
  - What a challenges do young people face in the job-market?
  - How did hiring firms view the certification from graduates
- Because for these questions we are less focused on giving a specific estimate of a population it is typical to use non-random sampling.

# Non-random sampling

---

- Non-random sampling means that the selection of the sample is not purely down to chance.
  - In some cases this may actually be desirable
  - Study and sample can be guided by findings
  - Can help if the sampling frame is not clear
- Does not necessarily mean the idea of an unbiased sample is always completely ignored – still important often to get a range of views.

## Sampling and CIE

# Credibility of a CIE

---

- Two main concerns when drawing conclusions from a CIE:
- Incorrectly conclude that the program has an impact when it doesn't in reality:
  - Consequence: Policy-makers believing the project does work when it doesn't and spending money on resources to expand it with no benefit to the population
- Incorrectly conclude that the program doesn't have an impact when it does in reality.
  - Consequence: Program is not-scaled up despite being beneficial to the population
- How can we help to avoid these?
- Ensure that the evaluation has a sufficient sample size

# Credibility of a CIE

---

- Set a high bar for confidence in the results
- Burden of proof: Before CIE, no impact is assumed. It is the task of the evaluator to prove otherwise, if there is an impact.

# Credibility of a CIE

---

- **Confidence level:** states the probability of your outcomes being accurate
  - Example 5% confidence level: We can say that there is a 5% probability of our outcomes being significant (or not) only by chance

**Larger  
Sample size**



**Higher  
Confidence  
Level**



**Lower chance of  
incorrectly concluding  
there is an impact**

# Credibility of a CIE

- **Statistical power:** Probability of finding a statistically significant difference between the groups, given that there is a difference in the population



**Larger  
Sample size**



**Higher  
Power**



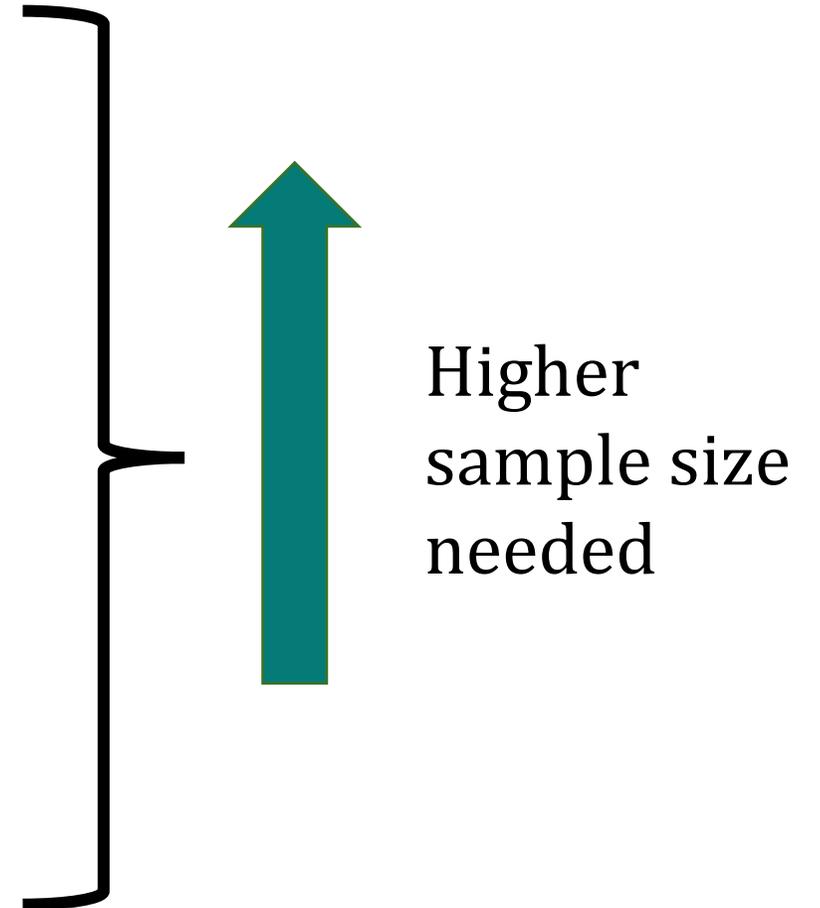
**Lower chance of  
incorrectly missing  
impact**

- To have a more in-depth understanding of how:
  - Sample size has an effect on the ability of an evaluation to detect impact and,
  - How it effects our ability conclude that the impact is not down to chance
- And how to actually calculate the required sample size for a credible CIE
- We would recommend this (not overly technical) lecture by Rachel Glennerster (Associate Professor of Economics at the University of Chicago)

[Sampling and Sample Size](#)

# Sample Size Determination

- Smaller the estimated effect of the program
- Lots of variation in the outcome of interest before the program
- Low take-up of the program
- High levels of attrition (losing participants over time)
- Estimating impact on sub-groups (such as males/females or host community/refugees)



# Sample Size and Non-Compliance

---

- Study power can be thought of as the strength of a magnifying glass.
- The smaller the thing you want to observe, the stronger the magnifying glass needs to be.
- Similarly the smaller the estimated effect of the program, the more power (and sample size) is needed.
- Non-compliance is when as selected participants do not take part in the program
  - e.g. Not using the Village Savings and Loan Association for any of their services



# Sample Size and Non-Compliance

- This dilutes the effect of the program as the effect for these people is expected to be zero
- This is obviously interesting for an evaluation of a policy – in reality people are going to always take up a program
- However for a CIE does not take into account when calculating the sample size it may make the results less credible
- May not have enough power and incorrectly conclude a program had no effect
- Improvement on implementation rather than the ToC



# Sample Size and Attrition

---

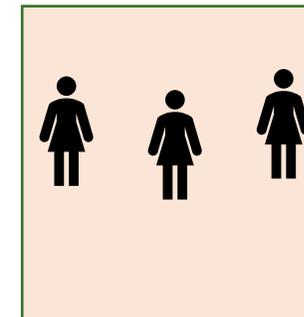
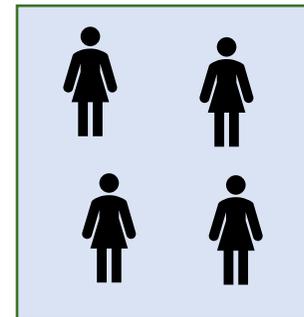
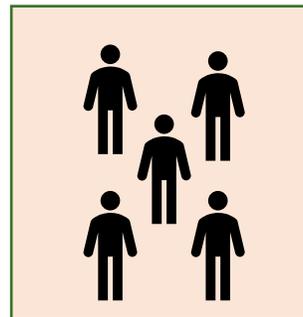
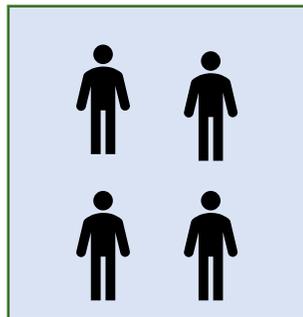
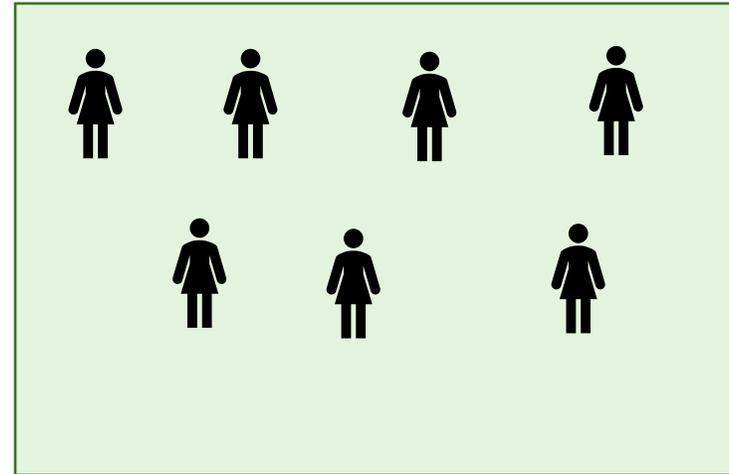
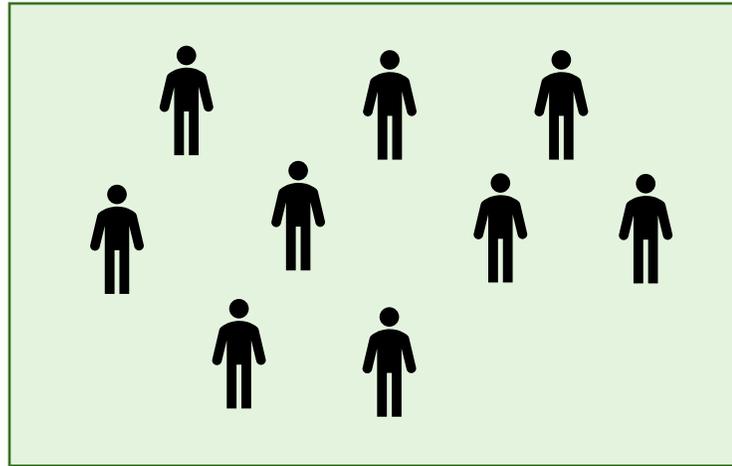
- Attrition is the loss of participants during a study
- Depends on:
  - Context -
  - Profile of participants – more mobile population
  - Implementation of study – fatigue
- Attrition is a big threat to a study in terms of bias
- However it is also an issue for sample size as non-response reduces the overall sample size
- Should be taken into account when calculating sample size

## Estimating Impact on sub-groups

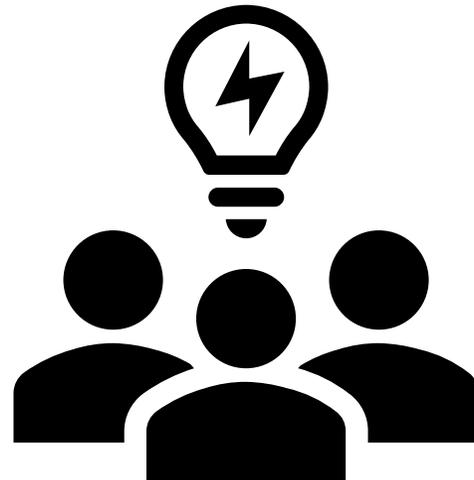
---

- If you have the research questions “What was the impact on females that took part in the vocational training program?”
- Then you have to consider this in the sampling strategy and sample size calculations.
- To estimate impacts for sub-groups (gender, refugee status, urban/rural) you need enough power for each of these sub-groups and ensure you have enough observations in each sub-group.
- Typically use stratified sampling for these groups

# Stratified Random Sampling



## Breakout Session 2



# Breakout Session

- **Scenario:** You are a team of evaluators for a youth vocational training program for refugees and host community members
- You have conducted a power calculation and found that to have a credible CIE, you need to have a total sample size of 2,500 respondents
- The endline study will take place 2 years after the end of the training
- The programme team have suggested, based on previous experience, that they expect a fair proportion of those randomly selected to participate in the program to not end up taking part
- Currently the budget only covers 2,000 respondents.
- **Task:** You should come up with a 2-minute presentation to the finance team to explain the benefits of increasing the total sample by 1,000 respondents. (Be prepared to answer the question why you should aim to increase above 2,500!). One team will be randomly selected to present to the rest of the teams.

# Sampling in challenging contexts

---

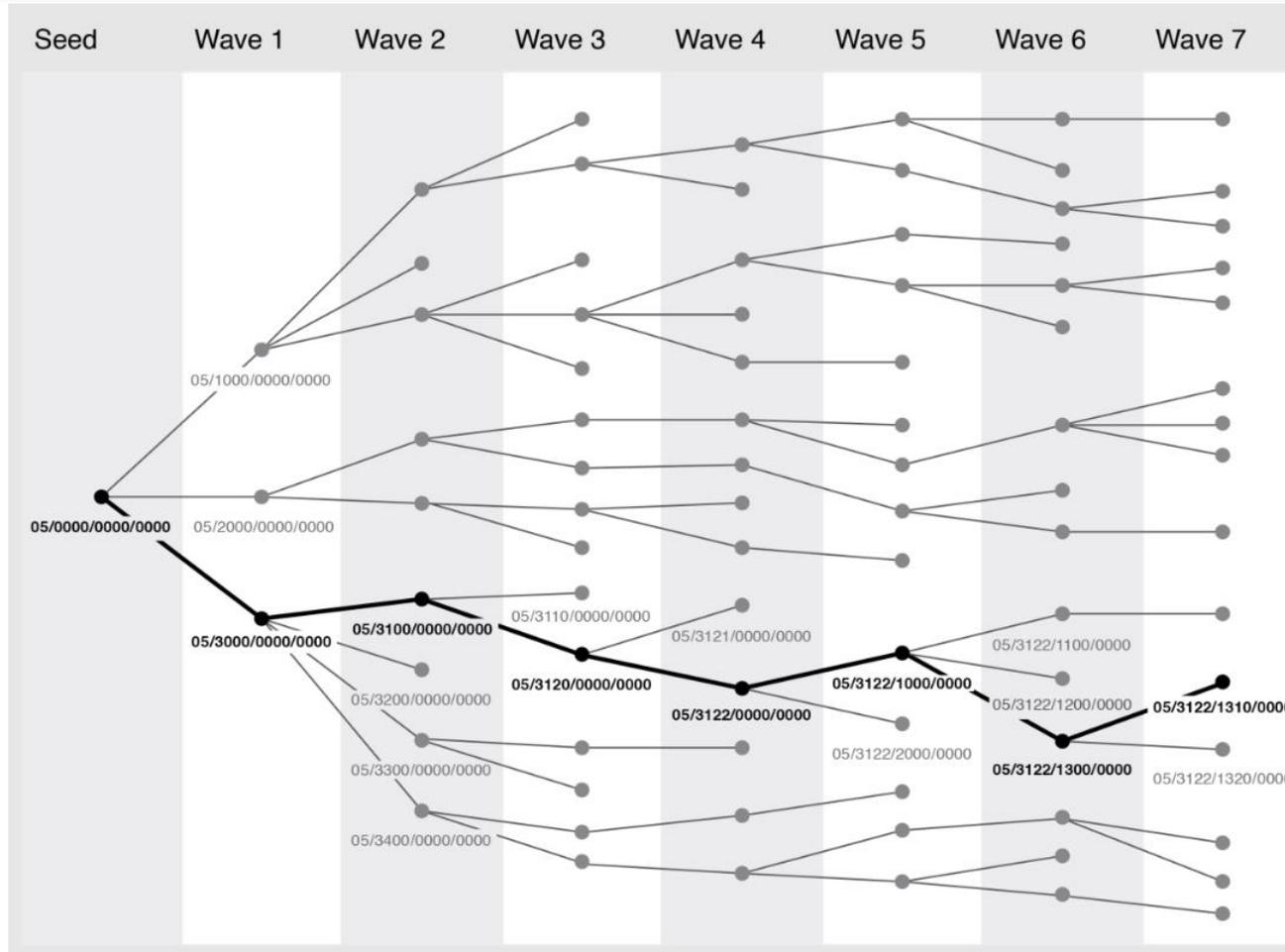
- In certain populations or settings it may be difficult to get sample frames
- Some populations are “hidden” because they belong to a socially stigmatized groups
  - Examples could include illegal migrants, HIV positive people and sex workers
- There is a lack of information about these populations and even if a researcher attempted to gather information on them, there may be little cooperation.

# Respondent-driven sampling

---

- One approach to dealing with this is **Respondent-driven sampling**
- In this cases initial participants - or “seeds” - are chosen (not at random – possibly identified by an organisation working with the target population)
- These participants list the initials of people in their network that are in the same population (i.e. other sex workers)
- They then recruit a pre-agreed number of peers who then in-turn recruit a set number of peers until chains are big enough to be independent of the initial seeds (Heckathorn, 1997).

# Respondent-driven sampling



Source: Hipp et al. (2019)

# Respondent-driven sampling

---

- This procedure captures the size of each participant's network makes it possible to calculate the probability of respondents within the network being selected
- Therefore provides a way to control the bias caused by differently sized personal networks.
- It also reduces non-response bias as it uses the trust of peers (and incentives)

# Respondent-driven sampling

---

- This provides a more robust method of sampling (McCreesh et al. 2012) for hidden populations for statistical inference than:
  - Snowball sampling (Just simply asking respondents to recommend others)
  - Convenience sampling (Surveying those which are easiest to survey)

## DAY 1 QUIZ

- Kindly look in the chat section for the link to the quiz
- Enter your name and use the code below if required
  - Code: **29600244**



# END OF DAY 1

- Gugerty, M. K., & Karlan, D. (2018). *The Goldilocks Challenge: Right-Fit Evidence for the Social Sector*. Oxford University Press.
- Groves, R. M., Fowler, F. J., Couper, M., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology*. Wiley.
- Iarossi, G. (2006). *The Power of Survey Design : A User's Guide for Managing Surveys, Interpreting Results, and Influencing Respondents*. The World Bank
- Desiere, S., & Joliffe, D. (2018). Land Productivity and Plot Size: Is Measurement Error Driving the Inverse Relationship?, *Journal of Development Economics*, 130, 84-98.
- Deininger, K., Carletto, C., Savastano, S., Muwonge, J. (2012). Can Diaries Help Improve Agricultural Production Statistics? Evidence from Uganda, *Journal of Development Economics*, 98, 42-50.
- Blair, G., Imai, K., Zhou, Y.-Y. (2015). Design and Analysis of the Randomized Response Technique, *Journal of the American Statistical Association*, 110, 1304-1319.
- Heckathorn, D. D. (1997). Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations, *Social Problems*, 44, 174-199.
- McCreesh, N., Frost, S. D. W., Seeley, J., Katongole, J., Tarsh, M. N., Ndunguse, R., Jichi, F., Lunel, N. L., Maher, D., Johnston, L. G., Sonnenberg, P., Copas, A. J., Hayes, R. J. And R. G. White (2012). Evaluation Of Respondent-Driven Sampling, *Epidemiology*, 23, 138-147.
- Hipp, L., Kohler, U., & Leumann, S. (2019). How to Implement Respondent-Driven Sampling in Practice: Insights from Surveying 24-Hour Migrant Home Care Workers. *Survey Methods: Insights from the Field*, 1-13