# Capacity Building Program for the Officers of Government of Meghalaya under P. A. Sangma Fellowship of Legal and Policy Research

*Module 2, 3, 4: Results based Framework, Impact Evaluation and Sampling Techniques*

# Project planning tools

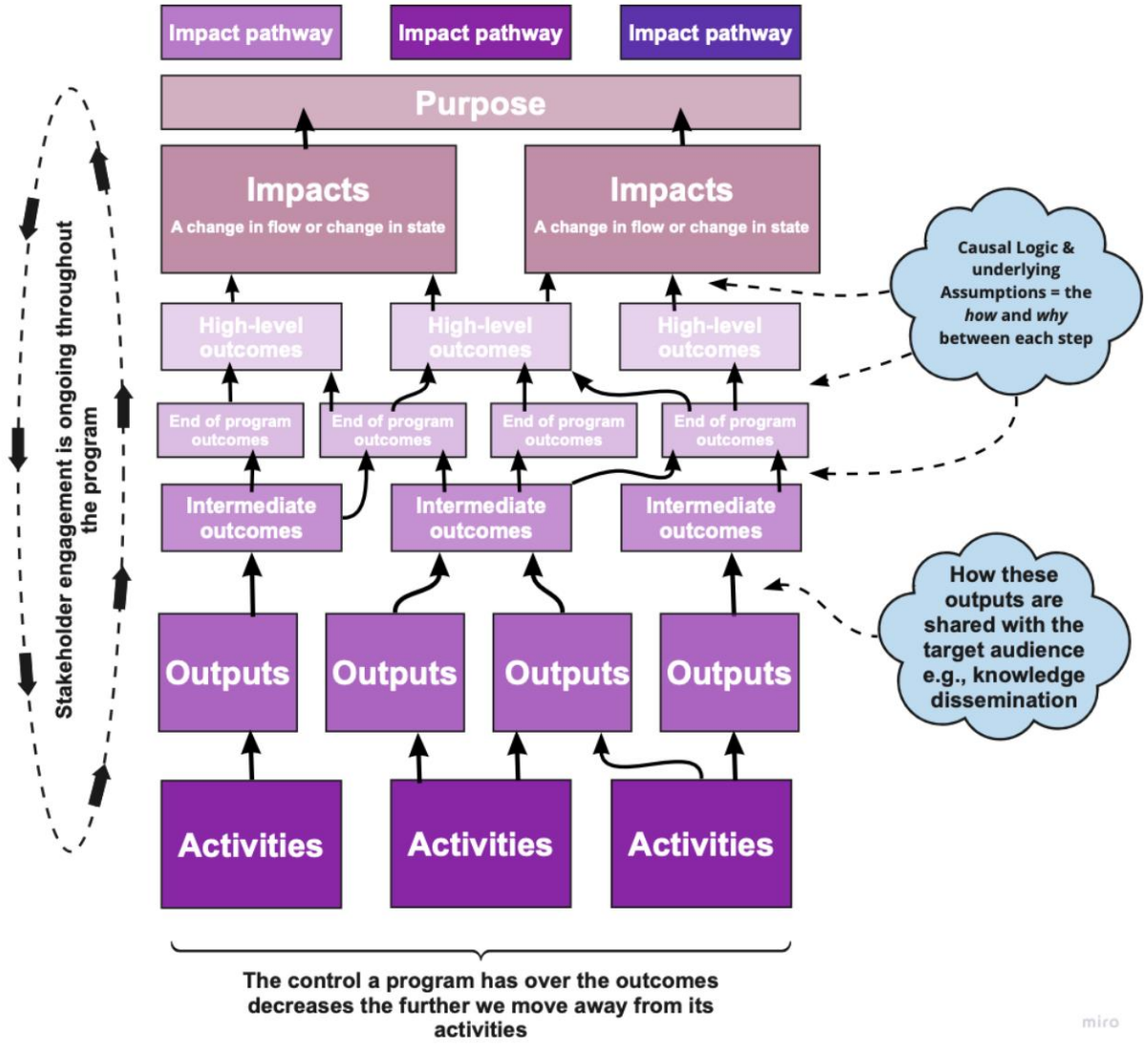
- Theory of Change (ToC)
- Log Frames
- Results Frameworks
- Logic Models

# Common points across tools

- All tools provide transparency and a visual explanation for why your program is expected to contribute to change
- All tools can help to track progress towards a specific objective
- All tools can be used at both the planning and evaluation stages of a program
- All tools are living documents and should be reviewed throughout the program lifespan
- All tools are time-consuming to develop, but to differing extents; they all require some reflexivity and strategic thinking to develop

# Theory of Change (ToC)

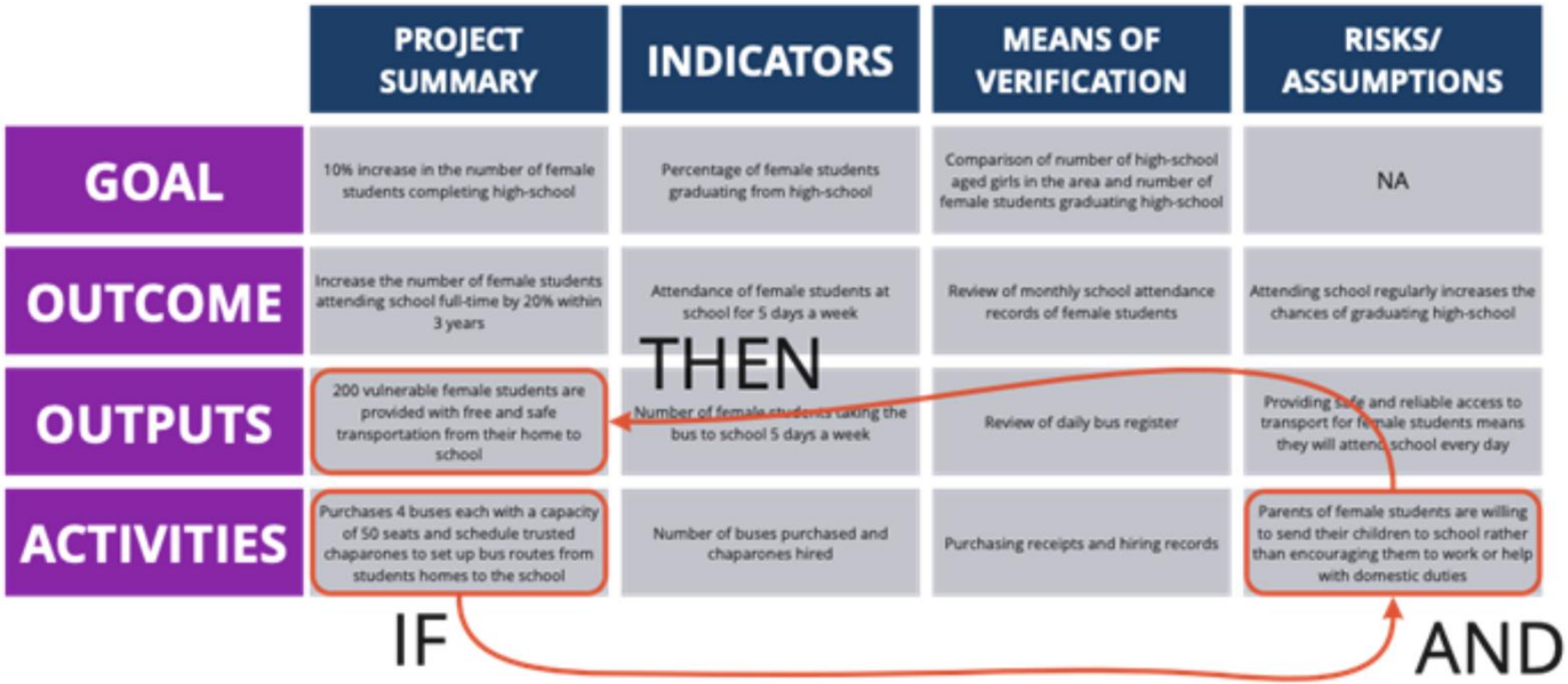
1. **What**
    - Explains how a set of activities will solve a problem through a diagram often made up of boxes and arrows
    - Goes into more detail by explaining the why also known as the “casual logic;” i.e., why one step is expected to lead to the next
2. **Why**
    - Explanatory and best suited to complex programs that are influenced by multiple systems
    - By defining long-term goals and then mapping backwards to identify necessary preconditions, ToC can provide the basis for arguing that a program is making a difference whilst identifying weaknesses in the argument and providing the opportunity to make changes
3. **When**
    - Can be used to both design and evaluate programs
    - Can be developed at any stage of an intervention
    - ToCs are living documents and should be flexible to the program’s needs and any changes happening on the ground
4. **Strengths**
    - ToCs capture unintended and unexpected results
    - Provides a reporting framework and identifies what data need to be gathered to test the theory
5. **Weaknesses -** Can be challenging and time-consuming as it involves facilitating collaboration with all key stakeholders, synthesizing a range of views and information sources, as well as obtaining agreement and buy-in from stakeholders


Impact pathway
Impact pathway
Impact pathway
Purpose
Impacts
A change in flow or change in state
Impacts
A change in flow or change in state
Causal Logic & underlying Assumptions = the *how* and *why* between each step
High-level outcomes
High-level outcomes
High-level outcomes
End of program outcomes
End of program outcomes
End of program outcomes
End of program outcomes
Intermediate outcomes
Intermediate outcomes
Intermediate outcomes
How these outputs are shared with the target audience e.g., knowledge dissemination
Outputs
Outputs
Outputs
Outputs
Activities
Activities
Activities
Stakeholder engagement is ongoing throughout the program
The control a program has over the outcomes decreases the further we move away from its activities

# Log Frames

1. **What**
    - Focused on how you will get to your program's goal
    - Usually presented as a matrix which structures the main activities in a program, highlights the logical connections between them, and identifies what these activities are expected to achieve
2. **Why**
    - Descriptive and better placed for small to medium sized projects
    - Log Frames help you to think about the relationships between available resources, planned activities, and the desired changes or results
3. **When**
    - Most Log Frames are developed during program design and are updated throughout the program's life span
    - Like ToCs, Log Frames are not set in stone and should be flexible to the program's needs
4. **Strengths**
    - It ensures objectives are clear and measurable
    - It ensures concrete evidence for a program's achievement is collected
    - Because risks and assumptions are made explicit, problems can be analyzed systematically
5. **Weaknesses**
    - It is a "one size fits all approach" which does not always capture the complexity and context of a program
    - Don't easily capture the how and why in the same way a ToC does

| | PROJECT SUMMARY | INDICATORS | MEANS OF VERIFICATION | RISKS/ ASSUMPTIONS |
|---|---|---|---|---|
| GOAL | 10% increase in the number of female students completing high-school | Percentage of female students graduating from high-school | Comparison of number of high-school aged girls in the area and number of female students graduating high-school | NA |
| OUTCOME | Increase the number of female students attending school full-time by 20% within 3 years | Attendance of female students at school for 5 days a week | Review of monthly school attendance records of female students | Attending school regularly increases the chances of graduating high-school |
| OUTPUTS | 200 vulnerable female students are provided with free and safe transportation from their home to school | Number of female students taking the bus to school 5 days a week | Review of daily bus register | Providing safe and reliable access to transport for female students means they will attend school every day |
| ACTIVITIES | Purchases 4 buses each with a capacity of 50 seats and schedule trusted chaparones to set up bus routes from students homes to the school | Number of buses purchased and chaparones hired | Purchasing receipts and hiring records | Parents of female students are willing to send their children to school rather than encouraging them to work or help with domestic duties |



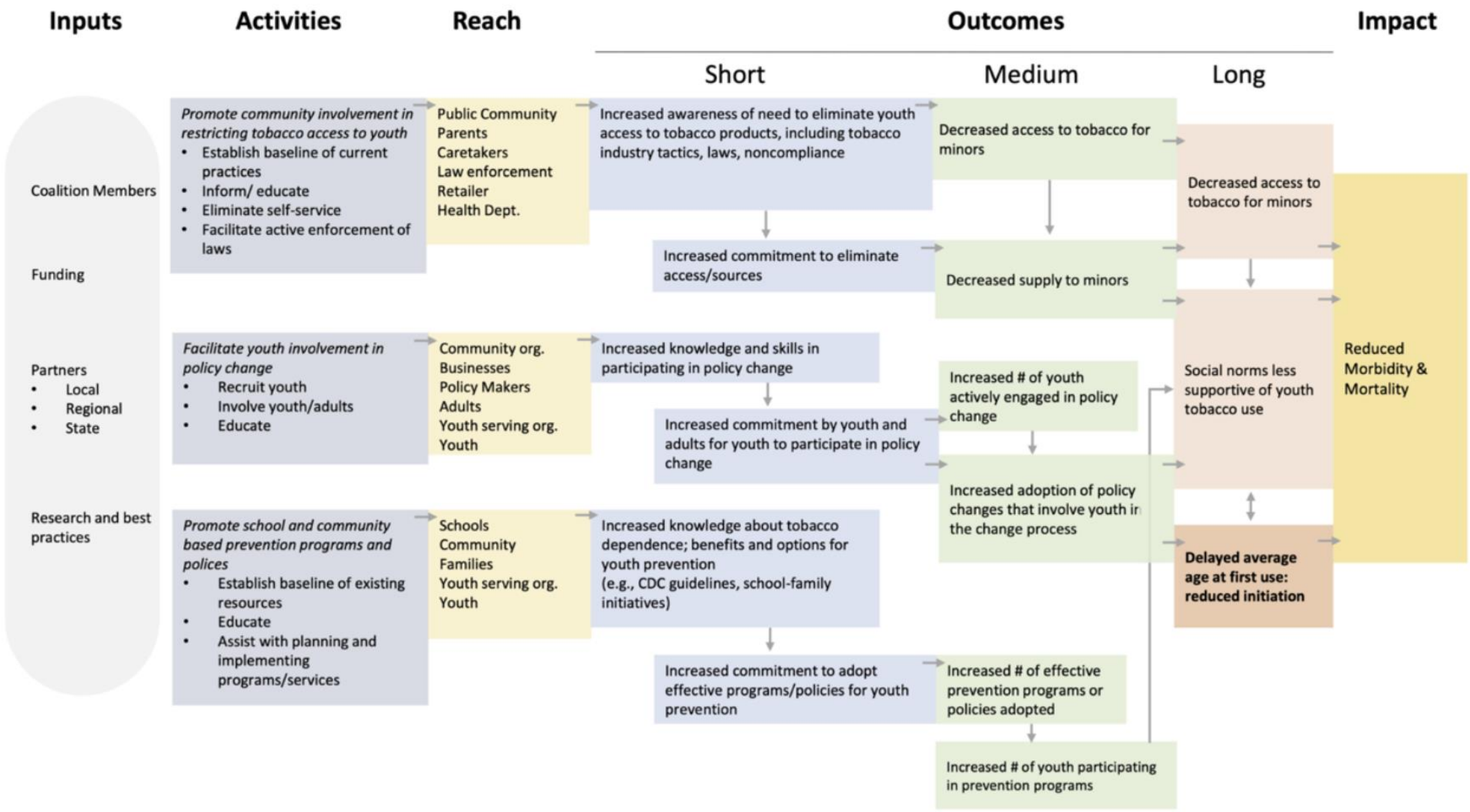
# Results Frameworks

1. **What**
    - Often in the form of a matrix that links activities with outcomes and results that directly relate to the objectives
    - Captures the essential steps of the logical and expected cause and effect relationship within a program
2. **Why**
    - Focus on explaining the program's results
    - It helps achieve strategic objectives i.e., the ultimate driver of the program by showing where resources could be best leveraged
3. **When**
    - Useful as part of a strategic planning process
    - Is a living management document to support consensus, guide course correction, and serve as an accountability framework for evaluation
4. **Strengths**
    - Helps identify and focus on specific, high leverage outcomes
    - Helps establish an evidence-based approach to monitoring and evaluation
    - Helps measure progress towards strategic objectives
5. **Weaknesses**
    - The effects of interventions can be difficult to fully measure as unintended consequences and external influences are not captured. This can lead to a risk of tunnel vision

| Activities | Outputs | Outcomes | Impacts |
|---|---|---|---|
| ***Main activities of the program*** | ***Expected main outputs of the program*** | ***A short- and medium-term effect of the project (usually not more than 2 or 3 per project)*** | ***Long-term effects produced by the program (usually 1 per program)*** |
| Act 1…<br>Act 2…<br>Act 3… | Output 1…<br>Output 2…<br>Output 3 …<br>Indicators…<br>Risks/assumptions… | Outcome 1…<br>Outcome 2…<br>Outcome 3…<br>Indicators…<br>Risks/assumptions… | Impact…<br>Indicators…<br>Risks/assumptions… |

# Logic Models

1. **What**
    - Usually presented in a flow chart (not a matrix)
    - Logic models visually summarise how a program is expected to work by listing: what resources will be used, what activities will be completed, and how the activities will lead to outcomes
2. **Why**
    - Logic models reveal intention, assumptions, and rationale behind a program
    - Logic models are useful to support stakeholders to think through and understand why a program is expected to lead to change
3. **When**
    - In the planning phase, logic models can help to shape program strategies, set priorities, and illustrate approaches to stakeholders
    - During program implementation, logic models can support accountability
4. **Strengths**
    - Builds a common understanding of goals, processes, and expectations for resources
    - Can help to explain the need for a program to the community, organization, or funder
    - Known for their easy-to-use format
5. **Weaknesses**
    - Don’t capture unintended or unexpected results
    - Don’t capture causality
    - While some logic models capture contextual factors and assumptions, they are often high-level and don’t look at each specific step within the change process


Inputs
Activities
Reach
Outcomes
Short
Medium
Long
Impact
Coalition Members
Funding
Partners
• Local
• Regional
• State
Research and best practices
Promote community involvement in restricting tobacco access to youth
• Establish baseline of current practices
• Inform/ educate
• Eliminate self-service
• Facilitate active enforcement of laws
Public Community
Parents
Caretakers
Law enforcement
Retailer
Health Dept.
Increased awareness of need to eliminate youth access to tobacco products, including tobacco industry tactics, laws, noncompliance
Decreased access to tobacco for minors
Decreased access to tobacco for minors
Increased commitment to eliminate access/sources
Decreased supply to minors
Facilitate youth involvement in policy change
• Recruit youth
• Involve youth/adults
• Educate
Community org.
Businesses
Policy Makers
Adults
Youth serving org.
Youth
Increased knowledge and skills in participating in policy change
Increased # of youth actively engaged in policy change
Increased commitment by youth and adults for youth to participate in policy change
Social norms less supportive of youth tobacco use
Increased adoption of policy changes that involve youth in the change process
Reduced Morbidity & Mortality
Promote school and community based prevention programs and polices
• Establish baseline of existing resources
• Educate
• Assist with planning and implementing programs/services
Schools
Community
Families
Youth serving org.
Youth
Increased knowledge about tobacco dependence; benefits and options for youth prevention (e.g., CDC guidelines, school-family initiatives)
Delayed average age at first use: reduced initiation
Increased commitment to adopt effective programs/policies for youth prevention
Increased # of effective prevention programs or policies adopted
Increased # of youth participating in prevention programs

# Results Framework

*A Practitioner's Approach*

# Why do we need it?

- It is difficult to know if programs have succeeded or failed if the expected results are not clearly articulated.
- An explicit definition of results—precisely what is to be achieved through the project or program and by when.
- Results-based management is a key tool for development effectiveness.

# How does a Results Framework help?

- A results framework serves as a key tool in the development landscape, enabling practitioners to discuss and establish strategic development objectives and then link interventions to intermediate outcomes and results that directly relate to those objectives.
- A results framework is an explicit articulation (graphic display, matrix, or summary) of the different levels, or chains, of results expected from a particular intervention—project, program, or development strategy.
- The results specified typically comprise the longer-term objectives (often referred to as “outcomes” or “impact”) and the intermediate outcomes and outputs that precede, and lead to, those desired longer-term objectives.

# Definition of key terms

- There are many debates, and considerable controversy, on the distinctions among outputs, outcomes, and impact.
- A generally useful approach is to consider outputs as the particular goods or services provided by an intervention (for example, nutrition supplements).
- An outcome is usefully thought of as benefits of that particular good or service to the target population (such as improved nutrition intake).
- Impact refers to evidence on whether outcomes are actually changing beneficiary behavior or longer-term conditions of interest (for example, improved eating habits, a healthier population).

# Pre-requisites to design a Results Framework

- An understanding of the problem or assessment of needs that the development intervention is intended to address.
- An initial theory of change for the project or program, even as it is being designed.
- A working knowledge of evidence required for measuring and assessing desired outcomes and impacts.
- Available data sources and proven data collection approaches relevant for the project or program context.
- These components provide a solid foundation on which to base a results framework.

# Steps in developing an effective Results Framework

- **Step 1. Establish Strategic Objective(s) for the Problem(s) to Be Addressed:**
  - A strategic objective is a calculated response to a known problem.
  - The wording and intent of the objective should be clear and specific enough that practitioners will be able to identify when it has been achieved.

# Tips for Effectively Stating Strategic Objectives

- Emphasize the results of actions, not the actions themselves. For example, instead of “reduce the transmission of HIV/AIDS,” use “reduced transmission of HIV/AIDS.” Instead of “promote credit opportunities for farmers,” use “increased credit available for farmers.”
- Maintain a single focus:
- Multiple objectives with multiple components are challenging to manage and measure.
- Test wording to avoid ambiguity:
- Test the wording with various stakeholders to ensure that the objective is consistently understood and not interpreted differently by different constituents.
- Specify the time frame:
- The amount of time available helps

- **Step 2. Identify and Work with Stakeholders**
  - In practice, various key parties (for example, government authorities or development partners) are already involved with the planning and/or implementation at this stage.
  - However, the team designing a results framework should revisit whether all the main stakeholders have been engaged to facilitate consensus and ownership of the initiative.
  - Wherever possible, the views and understanding of expected beneficiaries or target population should be considered in constructing the results framework.

- **Step 3. Define Results (Outputs and Outcomes)**
  - Outputs and outcomes represent those causal links in the results chain that bridge the gap between the current status and the desired high-level results. Starting with the end strategic objective(s), practitioners can backtrack to outline a program logic with immediate and intermediate outcomes.

- **Step 4. Identify Critical Assumptions and Risks**
  - Development interventions inevitably rely on some assumptions about factors that are beyond the control of the planners and implementers.
  - Results frameworks should not be based on critical assumptions that are perceived to have a low probability of holding true over the implementation period. If the risks are high, the intervention needs to be reconsidered.
  - For each output and outcome considered critical in the results chain, the framework developers should explicitly note assumptions related to external factors (for example, political environment, economy, climate change, and so forth) that could carry risks.
  - In cases where the assumption is seen to represent a more substantial risk, practitioners commonly adjust the development strategy, develop a contingency plan, and/or establish a risk management plan to monitor and address conditions as needed.

- **Step 5. Review Available Data Sources and Specify Indicators**
  - Where possible, measurement strategies should be based on existing data sources or tested data collection methodologies.
  - Relevant information for analysis and aggregation could already be available through administrative databases or through sample or census-based surveys.
  - Before specific indicators are defined for desired outcomes, practitioners should identify data sources that could be used to measure desired changes.

- **Step 6. Assign Indicators and Data Sources for Each Level of Result**
  - Strategic objective(s) and intermediate outcomes reflect constructs that need further definition to be measured. These outcomes need to be translated into a set of measurable indicators to establish whether progress is being achieved.
  - Indicators are tied to results by focusing on one or more characteristics of the outcome. A measure then expresses an indicator’s value quantitatively or qualitatively using SMART criteria

- **Characteristics of Effective—SMART—Indicators**
  - **Specific -** Indicators should reflect simple information that is communicable and easily understood.
  - **Measurable -** Are changes objectively verifiable?
    - Students' learning achievement
    - Value of land (number of hectares, multiplied by price per hectare)
    - Percentage of customers who are satisfied with the availability of potable water or electricity
  - **Achievable -** Indicators and their measurement units must be achievable and sensitive to change during the life of the project.
  - **Relevant -** Indicators should reflect information that is important and likely to be used for management or immediate analytical purposes.
  - **Time bound -** Progress can be tracked at a desired frequency for a set period of time.

- **Step 7. Establish the Performance Monitoring Plan**
  - The next step in designing a results framework is to plan how it will be operationalized to monitor progress and assess the effects of interventions. The plan for monitoring performance typically lists the following elements in a complementary tool, the monitoring plan:
    - Baseline and target values for selected measures to provide the means for verification to measure changes in the indicators
    - Data sources or methods for data collection.
    - The agent(s) responsible for collecting or providing the data (for example, independent evaluation team, project staff, and so forth).
    - Designated intervals at which the data will be collected or provided.
    - Assumptions and risks associated with the indictors or information being collected (such as the assumption that data will be available from a second party).

- **Step 8. Establish a Communication and Dissemination Plan**
  - The final step is to plan how the results framework will be used to communicate the progress and results of the intervention and how the results will be disseminated.
  - Some common approaches are to include results in a “dashboard,” highlighting only the key high-level objectives and outcomes/outputs achieved, using the framework for planning and review meetings (with the current status of the indicators highlighted), and using the change in the indicators from baseline to highlight the results.
  - Thus, choosing the correct outcome indicator (for example, change in rates of HIV) and connecting it to key intervention outputs (number of education campaigns about how HIV is transmitted) can provide a powerful communication and dissemination tool to inform and gather support from key stakeholders.

# Challenges

- Results frameworks offer clear benefits to practitioners and others working to achieve development results, but the approach is potentially challenging:
- An up-front investment of time and resources is needed at the start of an intervention.
- The effects of interventions can be difficult to measure fully.
- Results frameworks can become overly complicated. Attempts to apply
- Involving program staff in the evaluation process could bias results measurement.

# Evaluation Designs

*A Practitioner's Perspective*

# Contents

- What is a research design
- Why different research designs are needed?
- Causal – Experimental and Quasi-Experimental
- Exploratory Research – FGDs, Secondary Research, Qualitative Research, Expert Survey
- Descriptive – Cross-Sectional and Longitudinal
- Non-Experimental Design

# Learning Outcomes

- Research designs
- Basics of different evaluation designs
- Design experimental and quasi-experimental evaluations
- How to conduct descriptive and exploratory studies
- Know-how of non-experimental Designs
- Select evaluation design based on evaluation objectives, context, and resources

# Research Design

- Research design is the set of research methods and techniques chosen by a researcher to conduct a study.

- Characteristics of Research Design:
  - Neutrality: Result projected from the research should be free from bias (both external and internal)
  - Reliability: The research plan should indicate how the questions are formed to ensure the standard of results.
  - Validity: Correct measuring tools should be used to gauge results according to the objective of the research
  - Generalization: The outcome should apply to population and not just our expected sample

# Why is a design needed?

- Provides a firm foundation
- Reduces Uncertainty and confusion about the research problem
- Helpful for collecting research materials
- Gives an idea regarding the type of resources required in terms of money, human resources, time, and efforts
- Guides the research in the right direction
- And many more…….

# Research: Causal, Descriptive and Exploratory

## Causal

Conclusive Research- attempts to establish a cause-and effect relationship between two or more variables

Example: How much has COVID impacted the global economy

## Descriptive

Describe a phenomenon and its different characteristics – it describes the subject of the research without addressing why it happens

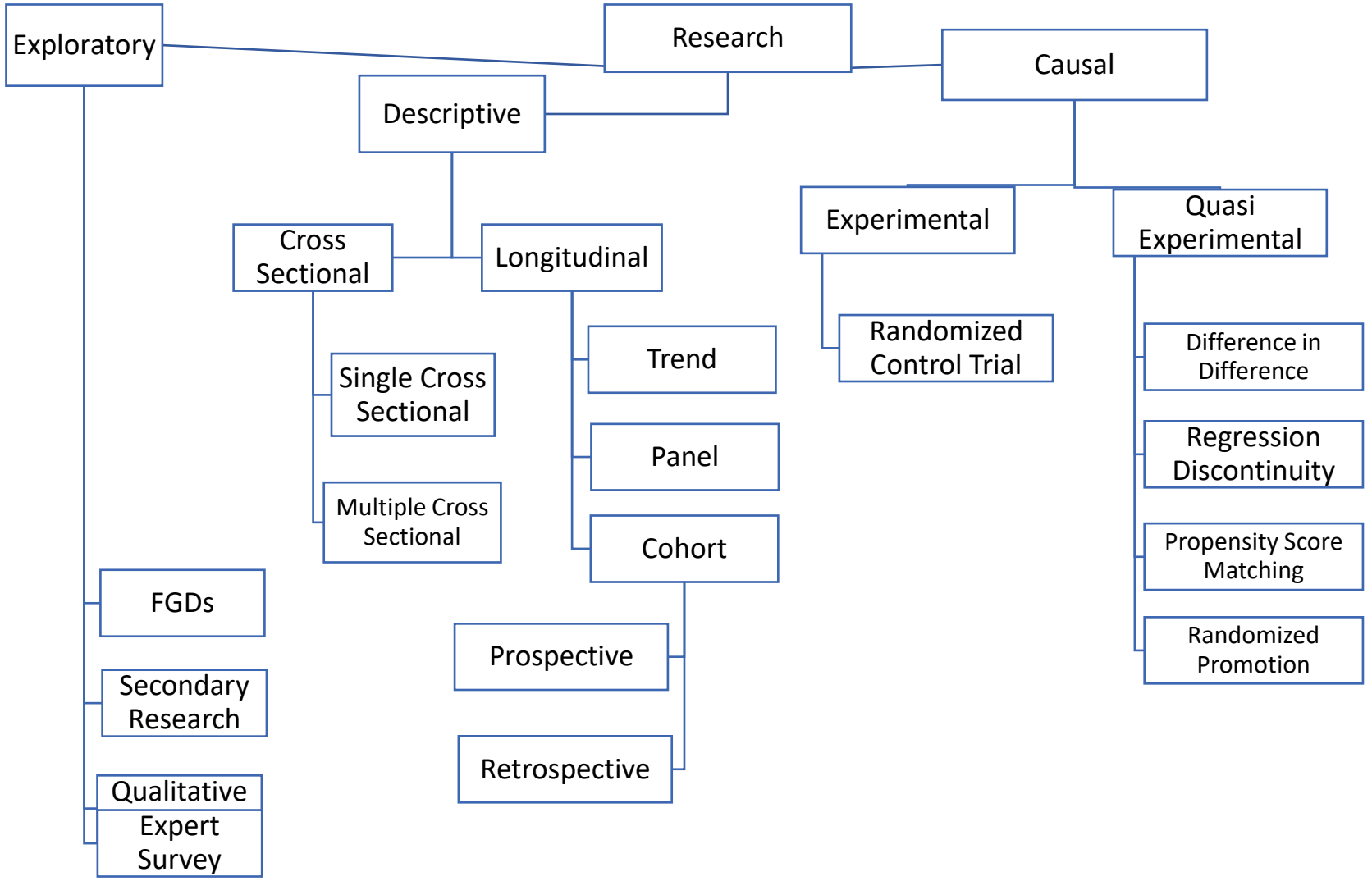
Example: Tracking changes in global economy post COVID.

## Exploratory

Investigate a research problem that is not clearly defined or understood – pave way for further research

Example: Study of Implications of COVID into the global economy

# Overview: Research Designs



# Causal: Experimental vs Quasi-Experimental

| | Experimental | Quasi-Experimental |
|---|---|---|
| Objective | Evaluate the effect of an intervention or a treatment | Evaluate the effect of an intervention or a treatment |
| Selection of participants | Random Assignment | Non-random assignment/Quasi-random assignment |
| Is there a control Group | Yes | Yes but not always |
| Is there any room for confounding | No | Yes |
| Level of evidence | Gold Standard of evidence | One below experimental design |
| Advantages | Minimum bias and confounding | Can be used in situations where an experiment is not ethically or practically feasible |
| Limitations | Cost, Ethical Considerations, Generalizability Issues, practically infeasible | Study is susceptible to bias and confounding, lower ranking in the hierarchy of evidence as compared to experimental |

# Impact Evaluation

- Objective - What is the **impact** (or causal effect) of a **program** on an outcome of interest?
- Central question -
  - **Magnitude** of change
  - **Attribution** of change
- Why IE?
  - Does programme achieve intended goal/s (measuring program efficacy)
  - Policy decision on whether to expand, modify or eliminate a prog.
- What would happen to the treatment group without the program?
- For attributing any change to the treatment i.e., establishing causality - a counterfactual or comparison group is required.
- Counterfactual
  - Group of nonparticipants(comparison group) which is statistically identical to the participant group in the absence of the program
  - program be ideally identical to the treatment group in observable and unobservable characteristics.
  - Treatment should  be the only difference between an ideal comparison and a project group
  - No other factor should influence the outcome apart from treatment

# Experimental Evaluations

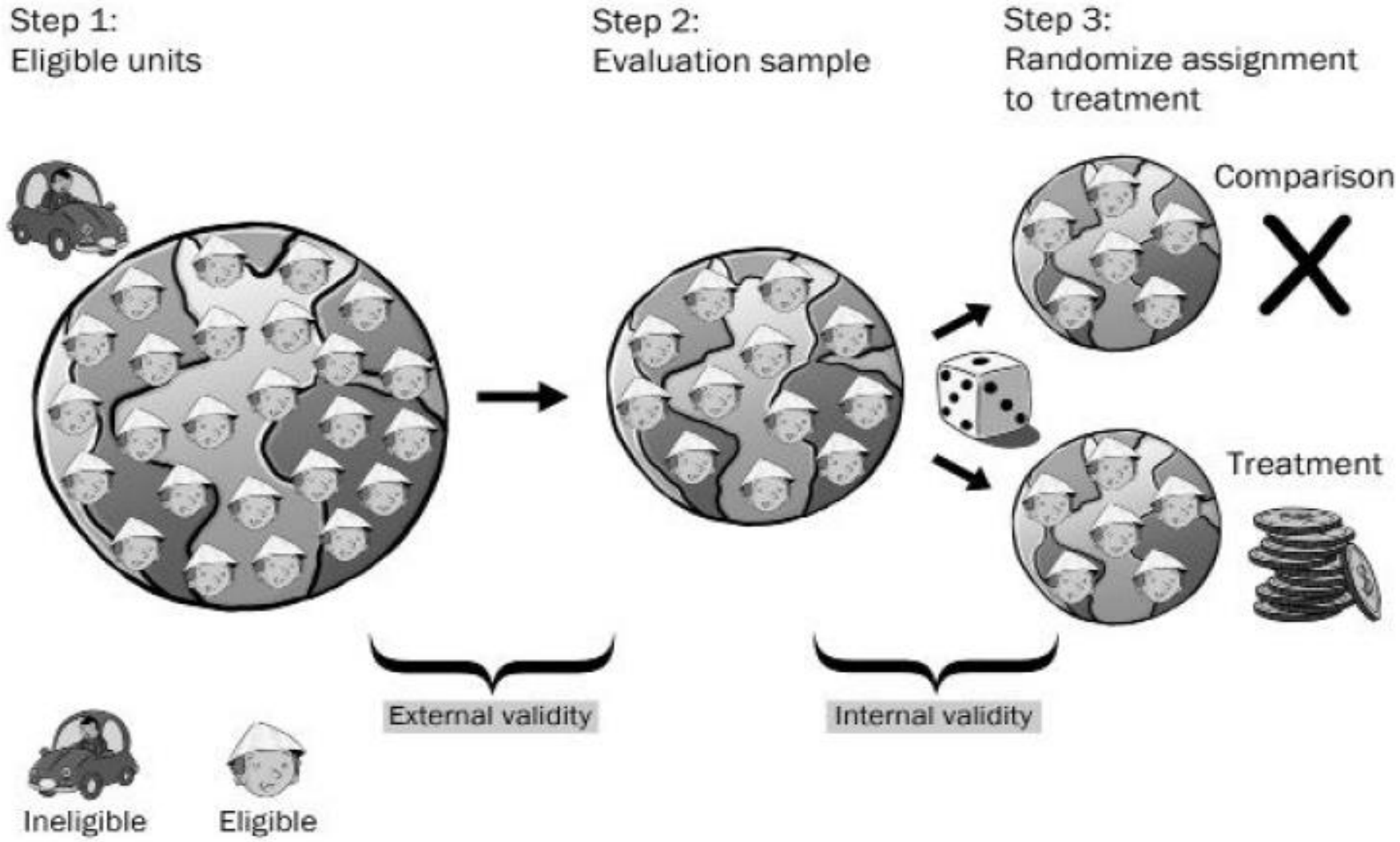
**Randomized Control Trial (RCT)**

Principle- The intervention is the only difference between the two groups

- This design involves gathering a set of equally eligible individuals willing to participate in the program and **randomly dividing** them **into treatment** and **control** groups.
- Most robust of all evaluation methodologies
- Control group generated through random assignment serve as the perfect counterfactual, free from selection bias.
- **Embedded** in the program **design phase**
- Randomization depends on the level at which the program is implemented
    - Individual level Randomized Control Trail(RCT)
    - Cluster RCT: Randomization done at the cluster level. e.g., at the village level/school level, etc.

# Experimental Evaluations: The Design

## Random Selection and Random Assignment



# Experimental Evaluation: Example

- To estimate the use of technology in increasing learning gains
  - A group of 200 students was randomly selected from a school and divided into two groups of 100 each.
  - One group was given a tablet with pre-loaded educational content, whereas the other wasn’t given anything
  - The learning levels of each group were measured after 1 year
  - The difference between project and control was the impact due to the intervention

# Experimental Evaluations

**Randomized Control Trial (RCT)**

Advantages

- Gold Standard of IE designs
- Closest counterfactual(identical group) is created
- Aims to remove or minimize any selection bias
- True measure of efficacy of programme

Limitations

- Ethical issues in implementation
- Political and Logistic Feasibility
- Expensive in terms of time and money
- Risk of contamination or spillover

# Experimental Evaluations

**Step Wedged/Pipeline Design**

- Used in special cases where program would be implemented in a phased manner
- Groups in which program would be implemented in later phases can be used as control
- Needs to ensure that the groups that are compared in phases are similar and comparable
- Appropriate gap should be between phases so that outcome can be achieved

# Step Wedged RCT: An Example

- To estimate the use of technology in increasing learning gains
  - Assume that the program will run for 3 years
  - A group of 300 students was randomly selected from a school and divided into three groups of 100 each.
  - 1st group was given a tablet with pre-loaded educational content in the first year, 2nd the group was given a tablet with content in the second year. Finally, the 3rd group was given the tablet with content in the 3rd year.
  - The learning levels of each group were measured each year.
  - In the first year, the difference between 1st group and 2nd (or 3rd) was the impact due to the intervention for 1 year
  - In the second year, the difference between the 1st group and 2nd group was the impact after 1 more year of the intervention. The difference between 1st and 3rd would be the impact of giving the technology for 2 continuous years.
  - This method would ensure that by the 3rd year all students get the technology.

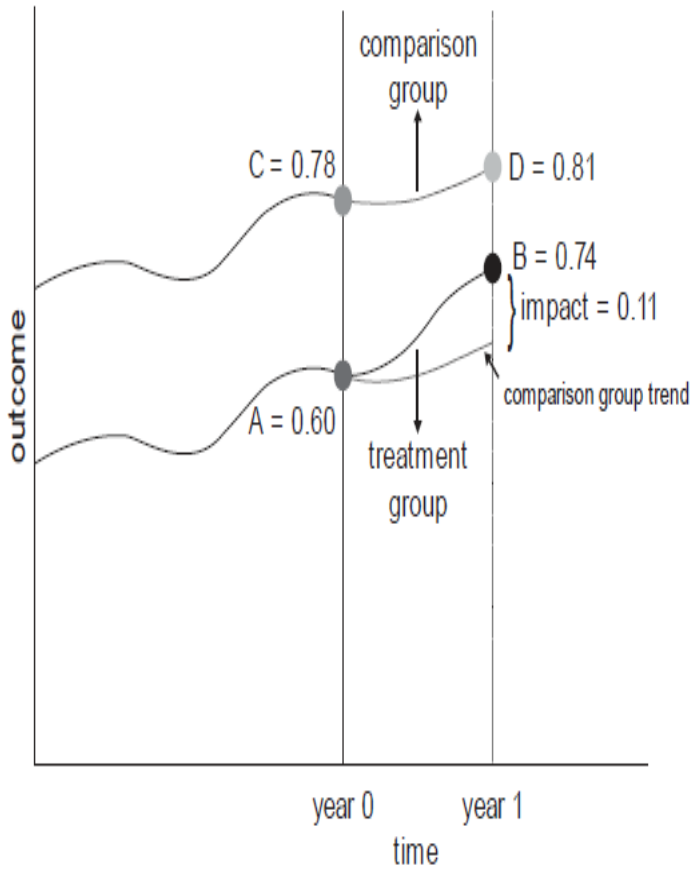
# Quasi Experimental

- Quasi or Semi experimental designs are not pure experimental designs i.e., **allocation** of **treatment** is **not random**.
- Are less robust than  experimental or randomized selection methods
- Control and treatment are not randomly distributed.
- Various techniques are used to create a counterfactual depending on the stage of program implementation and data availability

# Quasi Experimental - Difference in Difference(DID)

- DID  methods compare a treatment and a comparison group (first difference) before and after the intervention (second difference).
- This design requires at least two cross-sections of data, pre-program, and post-program, on treatment and control groups.
- DID controls the factors which are constant over time in both treatment and control
- Assumption: The selection bias is time-invariant (‘parallel trends’ in the absence of the program)

# Quasi Experimental - DID




| | After | Before | Difference |
|---|---|---|---|
| Treatment/enrolled | B | A | B − A |
| Comparison/ nonenrolled | D | C | D − C |
| Difference | B − D | A − C | DD = (B − A) − (D − C) |

| | After | Before | Difference |
|---|---|---|---|
| Treatment enrolled | 0.74 | 0.60 | 0.14 |
| Comparison/ nonenrolled | 0.81 | 0.78 | 0.03 |
| Difference | −0.07 | −0.18 | DD = 0.14 − 0.03 = 0.11 |

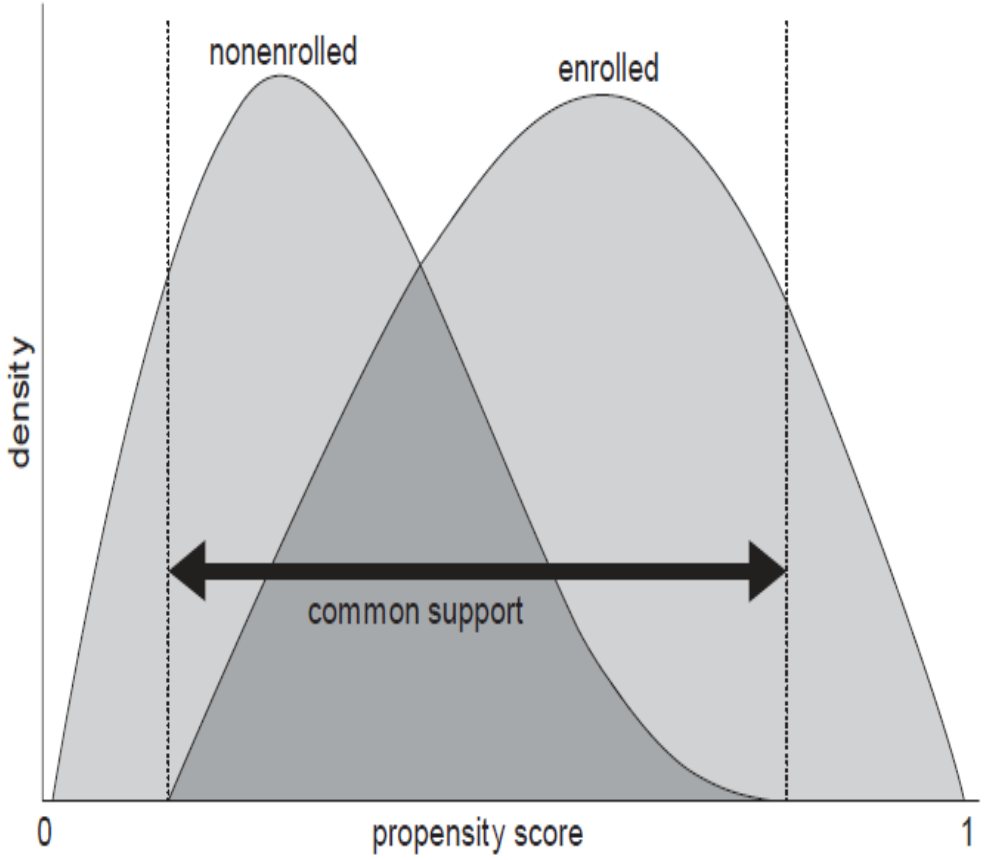
# DID: Example

- To estimate the use of technology in increasing learning gains
  - Let’s assume that instead of randomly selecting students, every student of one school (let’s say A) was given the content-loaded tablet.
  - In this case, you do not have a proper “counterfactual”
  - So, you randomly select some students for school A and some other school B (which you believe share some similar characteristics)
  - Establish a pre-post design in both case
  - DID approach would give to the impact of the intervention.

# Quasi-Experimental- Matching

**Steps in PSM**

1. Need representative and comparable data for both treatment and comparison groups
2. Propensity score or probability score of the two samples based on the selected observable characteristics needs to be calculated.
3. Restrict the sample to the units where common support appears in the propensity score distribution.
4. For each enrolled unit, locate a sub-group of non-enrolled units with similar propensity scores.
5. Compare the outcome of the treatment group with their matched comparison group. The difference in the average outcomes of the subgroups is the measure of the impact that can be attributed to the program.
6. Mean of these individual impacts yield the estimated average treatment effect.



# Example- PSM

**Project** : Impact evaluation of some Tribal Development Programme. Why PSM ?

- Retrospective Evaluation/ Ex post facto Evaluation
- Project ended in 2011
- No baseline data available
- How to attribute ?

Variables used to calculate the Propensity score to estimate likelihood of being involved in SHG

- Caste of the household
- APL/BPL card holders
- Literacy percentage
- Engagement in agricultural activities
- Participation in Gram Sabha meetings

# Example- PSM

Lets look at this fictitious scenario to understand the implications of PSM:

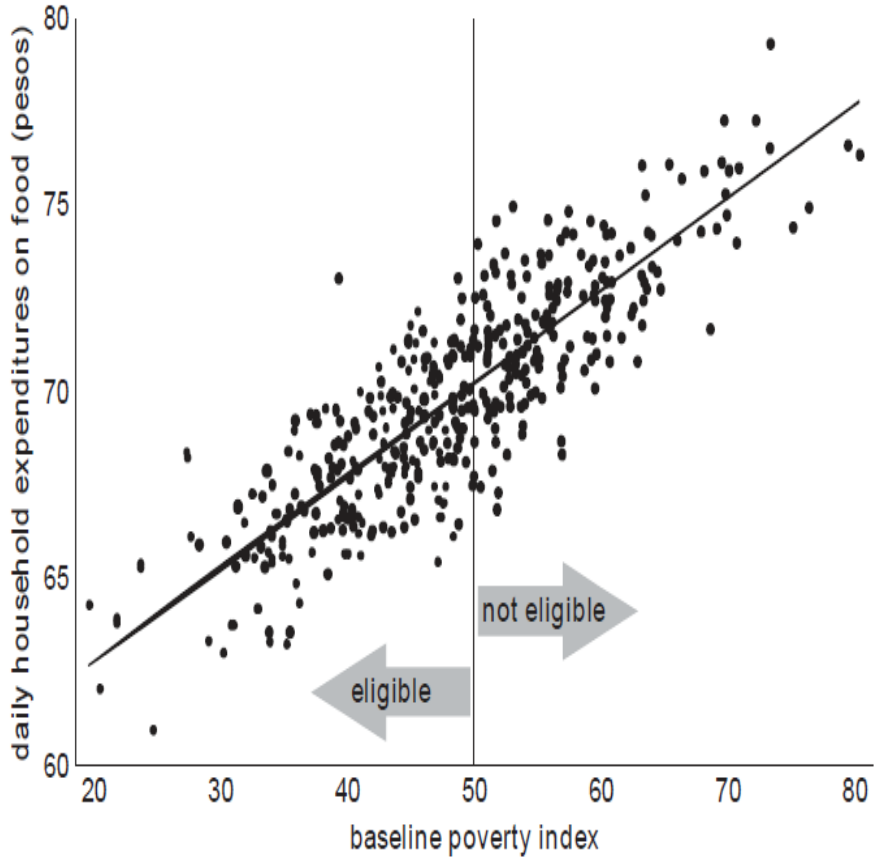
| | Before Matching | After Matching | Sample Loss |
|---|---|---|---|
| Project | 4448 | 3504 | 21.2 % |
| Control | 4356 | 3366 | 22.7% |

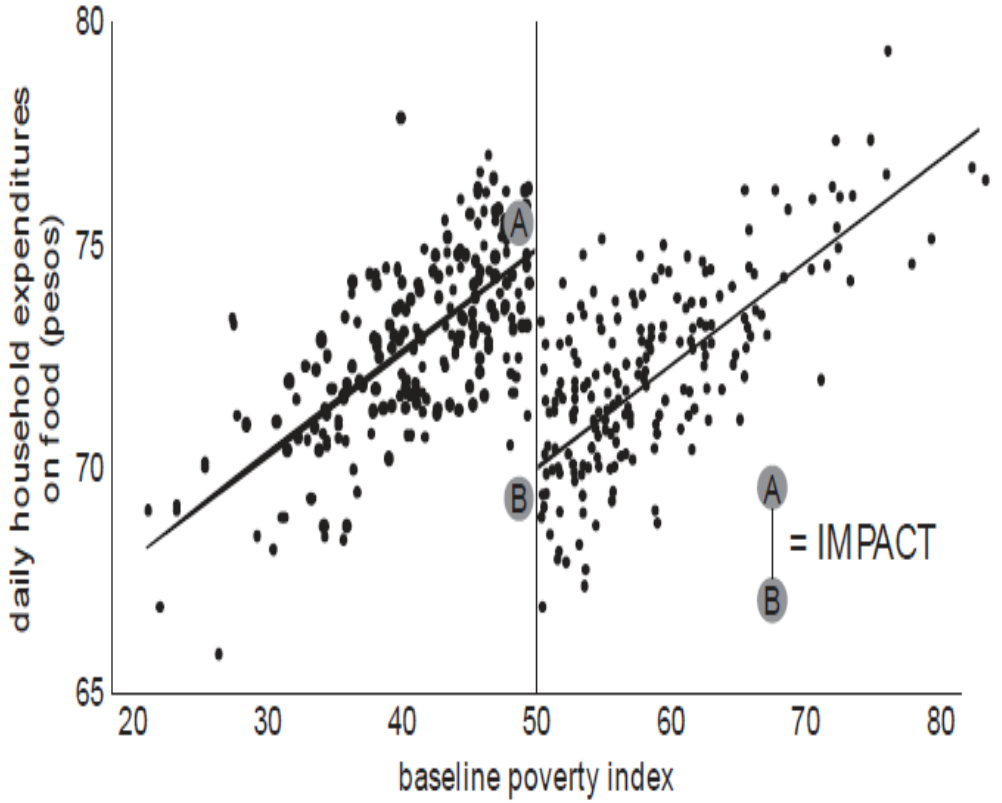
| Status | Indicator - Practice of regular Saving | Attribution |
|---|---|---|
| Before Matching | 4.5 % more households in treatment areas then in comparison areas | Intervention + confounding variables |
| After Matching | 2.4 % more households in treatment areas then in comparison | Minimizing the effects of confounding variables |

# Quasi Experimental - Regression Discontinuity

- RD can be used for programs that use a **continuous variable** or **index** to rank potential participants in a program.
- A cut-off point along the variable/index is used to determine the **eligibility criteria** i.e., whether or not a potential participant receives the program.
- E.g., Poverty score or index generated based on HH assets to deciding BPL households, pension programs based on retirement age cut-off.
- Candidates just below the cut-off point are similar to the candidates just above the cut-off point and hence can be treated as the control group.

# Quasi-Experimental - Regression Discontinuity



# RDD

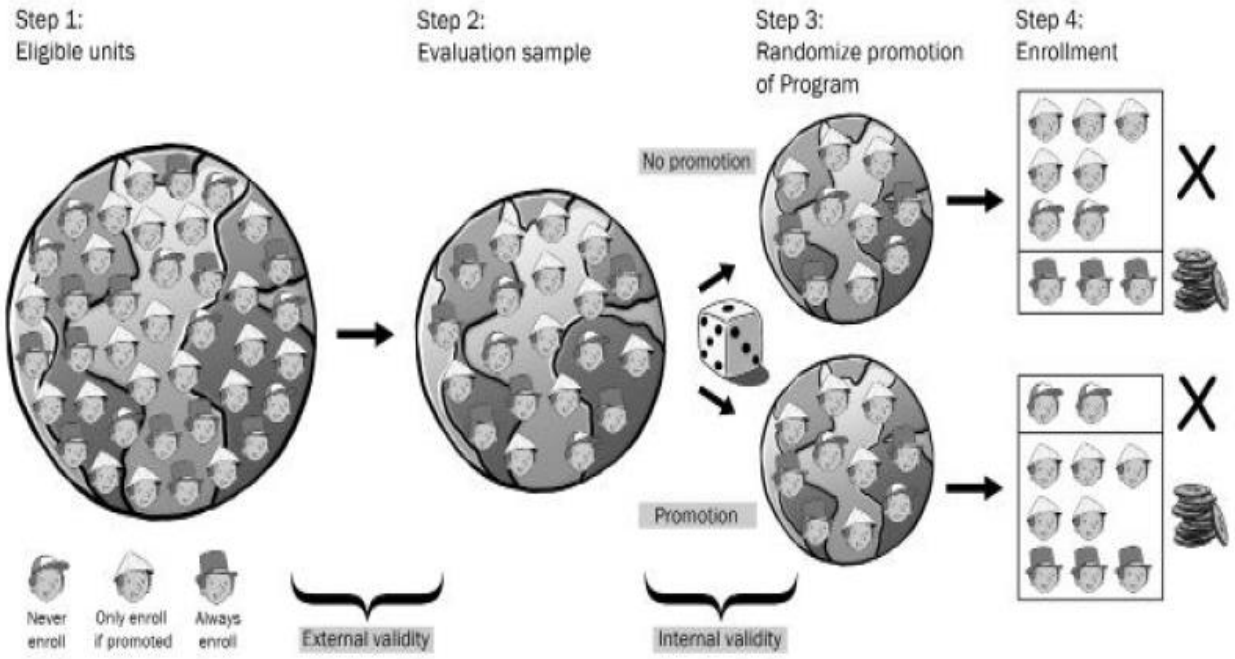
Regression discontinuity design (RDD) is an impact evaluation method that can be used for programs that have a **continuous eligibility index** with a **clearly defined eligibility threshold (cutoff score)** to determine who is eligible and who is not. To apply a regression discontinuity design, the following main conditions must be met:

1. The index must rank people or units in a continuous or “smooth” way.
2. The index must have a clearly defined cutoff score: that is, a point on the index above or below which the population is classified as eligible for the program.
3. The cutoff must be unique to the program of interest; that is, there should be no other programs, apart from the program to be evaluated, that uses the same cutoff score.
4. The score of a particular individual or unit cannot be manipulated by enumerators, potential beneficiaries, program administrators, or politicians.

# Quasi Experimental - Instrument Variable

- Applicable in case of programs having universal coverage and voluntary/open enrolment.
- Programme administrator cannot control who participates in the program. e.g., Skill development missions, and job training programs.
- Randomized promotion is an instrument variable that allows us to create variation between units and use this variation to create a valid comparison group.
- There are three types of participants. These are based on their intrinsic characters and cannot be measured by the program evaluator
  - Never Enroll
  - Always Enroll
  - Enroll if promoted
- Impact is measured based on the enrollment of promoted group
- Variable that affects the participation in the program, but not the program's outcome is identified.
- Counterfactual can be created by a random distribution of this variable in treatment and control groups. E.g. An outreach worker would randomly select the individuals to whom he/she visits and explains the program to.
- Assumption:
  - The promoted and non promote groups must be comparable.
  - Instrument variable should be effective enough to increase enrollment substantially amongst enrolled if promoted group.
  - Promotion activity should not affect the outcome of interest.

# Quasi Experimental- Instrument Variable




| | Promoted group | Non-promoted group | Impact |
|---|---|---|---|
| | % enrolled = 80%<br>Average Y for promoted group = 110 | % enrolled = 30%<br>Average Y for non-promoted group = 70 | Δ% enrolled = 50%<br>ΔY = 40<br>Impact = 40/50% = 80 |
| Never enroll | | | — |
| Only enroll if promoted | | | |
| Always enroll | | | — |

# Non- Experimental Designs

- NEDs are impact evaluation designs that do not include a matched comparison group
- Outcomes and impacts assessed without a conventional counterfactual to address the question
- Are used when
  - There are resource constraints: Used as default option when budget, time, data or other constraints do not permit the use of a “rigorous” evaluation design
  - Not possible to form a comparison group.
  - When program covers the entire population
- Limitation:
  - Cannot attribute the change as there is no counterfactual

# Non- Experimental Designs

## Some Potential NED

**Pre-test Post-test Design**

- Evaluators will survey the intervention group before and after the intervention
- Changes in outcome can be observed but cannot be attributed to the intervention

**Time Series Designs**

- Look for changes over time to determine the trend
- Evaluators observe the intervention multiple times before and after the intervention and analyze trends before and after

**Longitudinal Study**

- Another type of time series design
- Evaluated taking repeated measures of the same variable from the same people
- Panel design, a special type of longitudinal design.
- Small group of people is tracked at multiple points of time and record their experience in great detail.

# Non- Experimental Designs

## Ways to Strengthen Non-Experimental Design

- **Measur**e participants **level of exposure** to the programme: If participants with higher exposure shows higher change in outcome indicator, it strengthens the argument that programme lead to the change .
- **Collecting Data** from **same participants** over time using Longitudinal(Panel Design) : Characteristics of individuals can be controlled based

# Case Study 1

The government has decided to evaluate a three-year-long program that aims to measure the impact on the net annual household income due to their State-sponsored National Livelihoods project. You have been requested to propose a suitable evaluation design. Please propose a few options while listing the advantages and disadvantages of using each of the designs.

Ask these questions –

1. Do you have a Theory of Change, Results Framework, KPI matrix, or LFA?
2. When are you onboarded? Before the start of the project or after or midway?
3. What do you want to measure?
4. In which areas?
5. Do you want estimates at the project-level or national-level, or sub-division level?
6. If possible, what is the overall resource availability like for the evaluation?

Read about the program. Talk to the government. If possible, conduct an exposure visit/s. Think about the problem statement before the design.

# Case Study 2

The Amazing Foundation has developed an innovative AI-powered technological solution to address the issue of pests in cotton cultivation. It helps the farmers detect an infestation, area of infestation, type of pesticide required, and amount of pesticide, with video tutorials on the app, which can be used in offline mode also. It also helps get free online consultations with experts. This helps reduce farmers' input costs and time, and prevents crop loss. You are required to propose an evaluation design to check the effectiveness of the mobile app and its effect on the yield. Please suggest a design while mentioning the reasons for not adopting the other designs.

# Case Study 3

In the absence of a baseline, please design an evaluation of a 5-year long program aimed at improving the overall well-being of the households –

- Physical well-being.
- Economic well-being.
- Social well-being.
- Development and activity.
- Emotional well-being.
- Psychological well-being.
- Life satisfaction.
- Domain-specific satisfaction.

# Case Study 4

A collective of farmers that engages in producing and collectively selling the produce in the market is called Farmer Producer Organization. This organization is registered under some National Act. Please design an evaluation to assess the performance of these FPOs in the last three years.

# Case Study 4

A campaign was launched in certain geographies to improve awareness on climate change and clean air. For this, the AQI index was displayed on large boards with certain key messages on damage to human health and the environment due to climate change and ways to contain/minimize it. Please design an evaluation to measure the impact of this program.

# Case Study 5

A long-term government program is being implemented in three phases to improve the WASH status of the households.

- Phase 1 – Community mobilization between 2023-2024
- Phase 2 – Implementation phase between 2024-2026
- Phase 3 – End of project phase (project withdrawal phase) 2026-2027

Please design an evaluation to measure the impacts of the program

# Case Study 6

A long-term government program is being implemented in three phases to improve the learning outcomes of students in grade 5 using adaptive learning platforms.

- Phase 1 – 50 schools will be included
- Phase 2 – 60 schools will be included (110 schools in total)
- Phase 3 – 80 schools will be included (190 schools in total)
- Phase 4- 100 schools will be included (290 schools in total)
- Phase 5- 200 schools will be included (490 schools in total)

Please design an evaluation to measure the change in the learning outcomes of the students.

# Case Study 7

A food distribution policy has been launched by the government in a national emergency situation. Households have been assigned an index based on their economic situation and distance from the food distribution centre. It is a 3 – point index. Category 1 will be provided a 100% subsidy on food grains, category 2 with 50% and category 3 will be provided an 20% subsidy.

Propose a design to measure the impact of providing subsidy on food consumption scores of households.