An Outside Evaluator's Perspectives on the DCED Standard

Jim Rugh

OECD/DAC Definition of Evaluation

- The systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision– making process of both recipients and donors.

- Evaluation also refers to the process of determining the worth or significance of an activity, policy or program. An assessment, as systematic and objective as possible, of a planned, on-going, or completed development intervention.

There are a great many purposes and types of evaluation, and an infinite number of *evaluands* (the objects of evaluation).

- **Formative**: learning and improvement including early identification of possible problems

- **Knowledge generating**: identify cause-effect correlations and generic principles about effectiveness.

- **Accountability**: to demonstrate that resources are used efficiently to attain desired results

- **Summative judgment**: to determine value and future of program

- **Developmental evaluation**: adaptation in complex, emergent and dynamic conditions

Caution: Too often what is thought of as (or experienced as) External Evaluation is based on a "we will examine and judge you" paradigm.
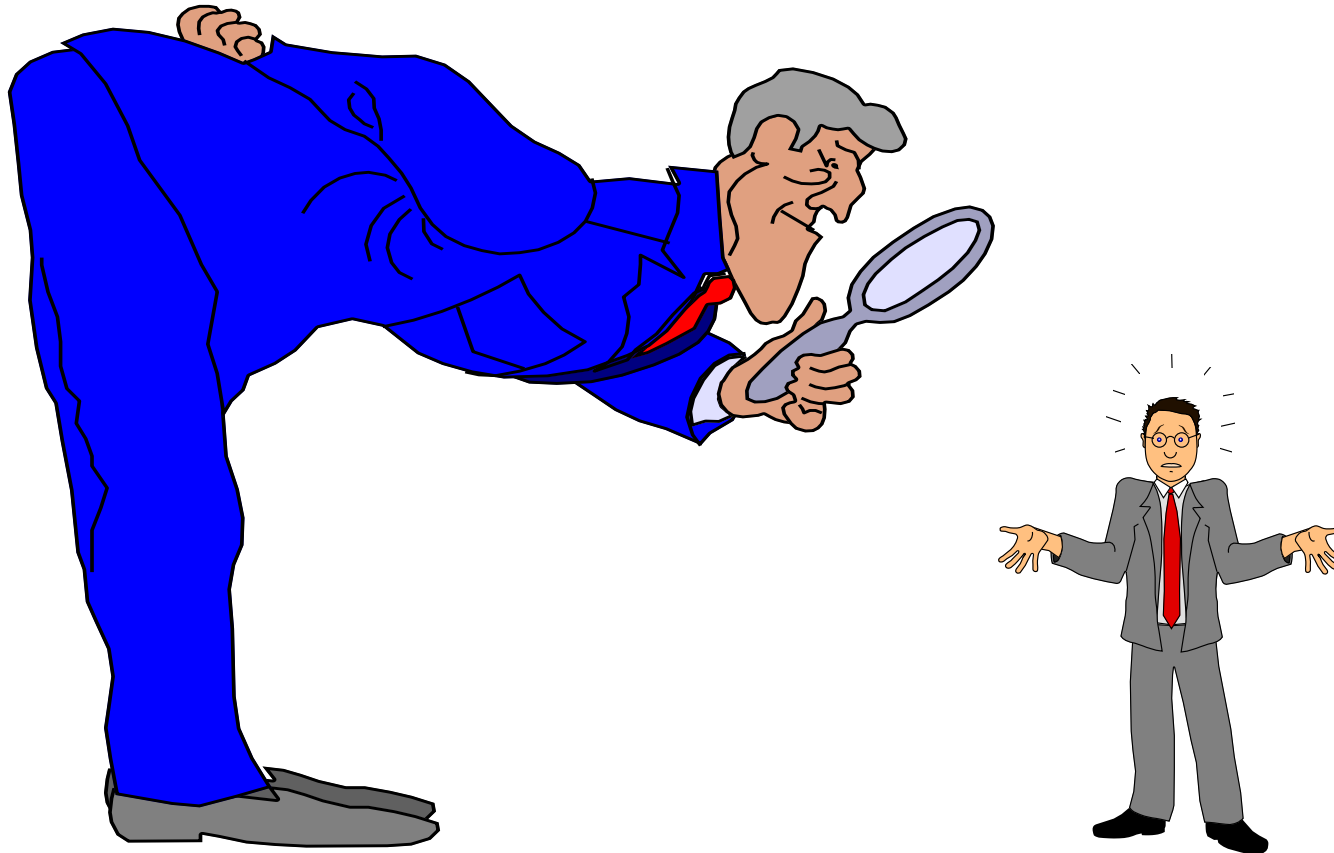
How much more helpful it is when the approach to evaluation is more like holding up a mirror to help people reflect on their own reality: *facilitated self-evaluation*.
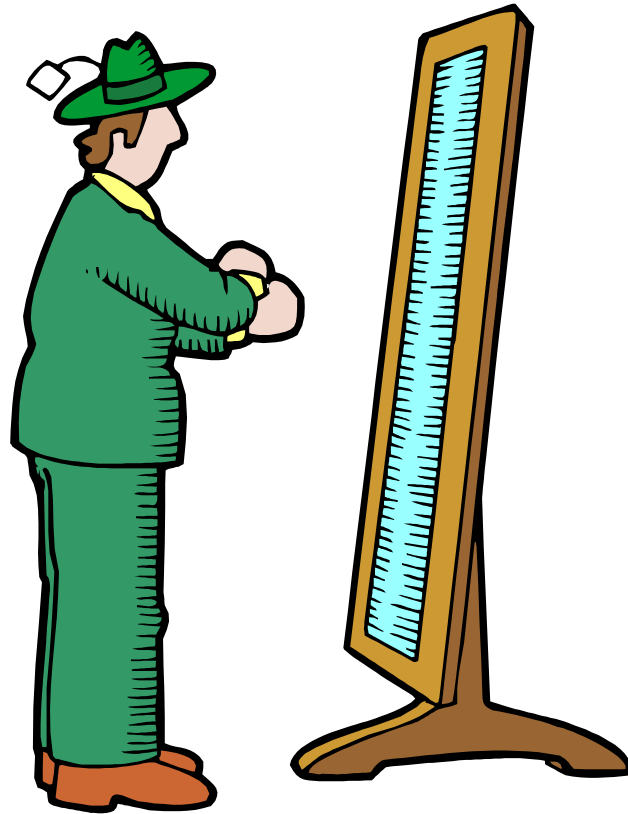
The eight elements of the DCED Standard relate directly or indirectly to what is commonly referred to as evaluation

1. **Articulating the Results Chain** [Certainly the first step in the evaluation process, at least getting clarity on what the program planners had in mind.]

2. **Defining indicators of change** [Obviously necessary for evaluation, preferably at outcome or impact levels.]

3. **Measuring changes in indicators** [Commonly the monitoring process looks more at indicators at the level of activities and outputs; evaluation is often required for measuring outcomes and impact.]

4. **Estimating attributable changes** [Refers back to the Results Chain plus looking for plausible counterfactuals. An important evaluation function.]

The eight elements of the DCED Standard relate directly or indirectly to what is commonly referred to as evaluation

5. **Capturing wider changes in the system or market**. [Both M & E should look for what other influences there were, internally or externally.]

6. **Tracking programme costs**. [Normally a financial accounting and auditing function.]

7. **Reporting results** [Obviously the monitoring system should generate frequent reports during the life of a program, but often it is the more extensive purview of a mid-term or end-of-project or ex-post evaluation that can generate more substantive reporting of verified results.]

8. **Managing the system for results measurement** [Obviously necessary for a good M&E system, from beginning to end.]

There are now such organizations in 111 countries.

That interactive world map and constantly updated database of Voluntary Organizations for Professional Evaluation (VOPEs) is accessible at

www.IOCE.net

RealWorld Evaluation

Designing Evaluations under Budget, Time, Data and Political Constraints

Jim Rugh

Note: The summary chapter of the book and other resources are available at:

www.RealWorldEvaluation.org

Your plan
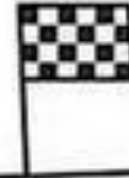
The RealWorld Evaluation Approach

An integrated approach to ensure acceptable standards of methodological rigor while operating under real-world budget, time, data and political constraints.

*See the RealWorld Evaluation book
or at least condensed summary for more details*

2
EDITION

RealWorld Evaluation

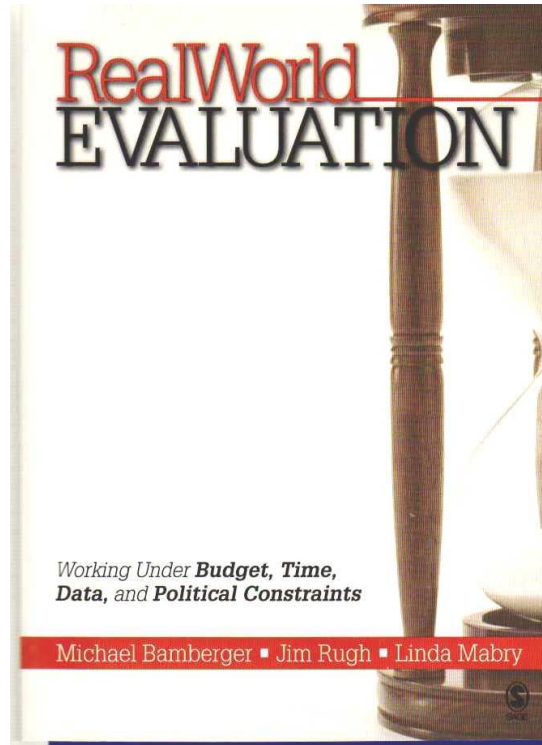This book addresses the challenges of conducting program evaluations in real-world contexts where evaluators and their clients face budget and time constraints and where critical data may be missing. The book is organized around a seven-step model developed by the authors, which has been tested and refined in workshops and in practice. Vignettes and case studies—representing evaluations from a variety of geographic regions and sectors—demonstrate adaptive possibilities for small projects with budgets of a few thousand dollars to large-scale, long-term evaluations of complex programs. The text incorporates quantitative, qualitative, and mixed-method designs and this Second Edition reflects important developments in the field over the last five years.

New to the Second Edition:

- **Adds two new chapters on organizing and managing evaluations,** including how to strengthen capacity and promote the institutionalization of evaluation systems

- **Includes a new chapter on the evaluation of complex development interventions,** with a number of promising new approaches presented

- **Incorporates new material,** including on ethical standards, debates over the "best" evaluation designs and how to assess their validity, and the importance of understanding settings

- **Expands the discussion of program theory,** incorporating theory of change, contextual and process analysis, multi-level logic models, using competing theories, and trajectory analysis

- **Provides case studies of each of the 19 evaluation designs,** showing how they have been applied in the field

"This book represents a significant achievement. The authors have succeeded in creating a book that can be used in a wide variety of locations and by a large community of evaluation practitioners."

—Michael D. Niles, *Missouri Western State University*

"This book is exceptional and unique in the way that it combines foundational knowledge from social sciences with theory and methods that are specific to evaluation."

—Gary Miron, *Western Michigan University*

"The book represents a very good and timely contribution worth having on an evaluator's shelf, especially if you work in the international development arena."

—Thomaz Chianca, *independent evaluation consultant, Rio de Janeiro, Brazil*

RealWorld Evaluation

2
•EDITION

RealWorld Evaluation

Working Under Budget, Time, Data, and Political Constraints

Michael Bamberger **Jim Rugh** **Linda Mabry**

2
EDITION

SAGE

What is special about the RealWorld Evaluation approach?

- There is a series of steps, each with checklists for identifying constraints and determining how to address them
- These steps are summarized on the following slide …

The Steps of the RealWorld Evaluation Approach

Step 1: Planning and **scoping** the evaluation

Step 2: Addressing **budget** constraints

Step 3: Addressing **time** constraints

Step 4: Addressing **data** constraints

Step 5: Addressing **political** constraints

Step 6: **Assessing** and **addressing** the strengths and weaknesses of the evaluation design

Step 7: Helping clients **use** the evaluation

Step 1: Planning and scoping the evaluation

A. Defining client information needs and understanding the political context

B. Defining the program theory model

C. Identifying time, budget, data and political constraints to be addressed by the RWE

D. Selecting the design that best addresses client needs within the RWE constraints

Step 2
Addressing **budget** constraints

A. Modify evaluation design

B. Rationalize data needs

C. Look for reliable secondary data

D. Revise sample design

E. Economical data collection methods

Step 3
Addressing <u>time</u> constraints

All Step 2 tools plus:

F. Commissioning preparatory studies

G. Hire more resource persons

H. Revising format of project records to include critical data for impact analysis.

I. Modern data collection and analysis technology

Step 4
Addressing <u>data</u> constraints

A. Reconstructing baseline data

B. Recreating comparison groups

C. Working with non-equivalent comparison groups

D. Collecting data on sensitive topics or from difficult to reach groups

E. Multiple methods

Step 5
Addressing <u>political</u> influences

A. Accommodating pressures from funding agencies or clients on evaluation design.

B. Addressing stakeholder methodological preferences.

C. Recognizing influence of professional research paradigms.

Step 6
Assessing and addressing the strengths and weaknesses of the evaluation design

An integrated checklist for multi-method designs
A. Objectivity/confirmability
B. Replicability/dependability
C. Internal validity/credibility/authenticity
D. External validity/transferability/fittingness

Step 7

Helping clients use the evaluation

A. Utilization
B. Application
C. Orientation
D. Action

RealWorld Evaluation

Designing Evaluations under Budget, Time, Data and Political Constraints

LOGIC MODELS / THEORIES OF CHANGE / RESULTS CHAINS

"I think you should be more explicit here in step two."

An assumed (simple) theory of change (logic model)

Children are educated.

Parents send their children (including girls) to school.

Schools are built in or near villages where there were no schools before.
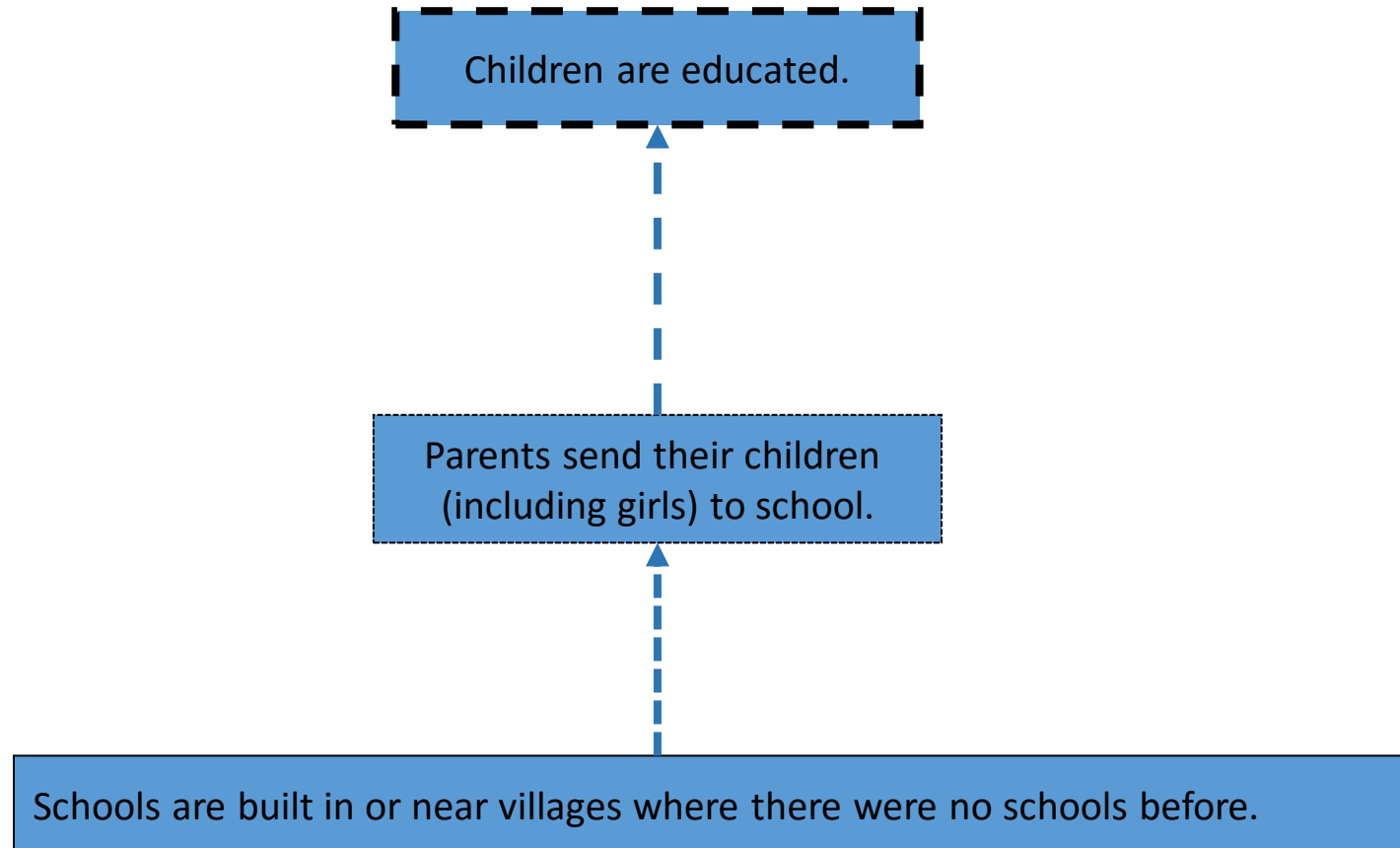
A (slightly) more comprehensive theory of change
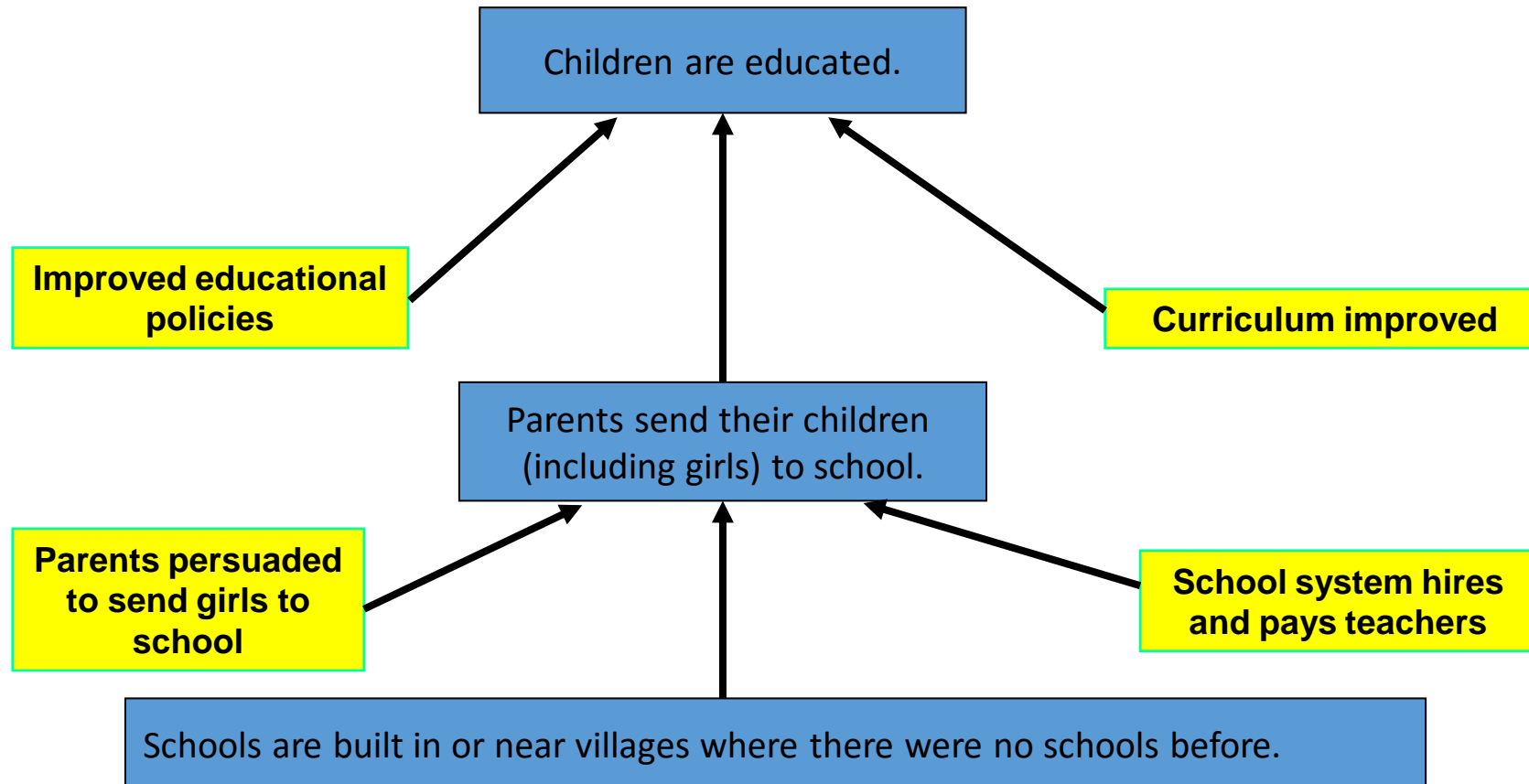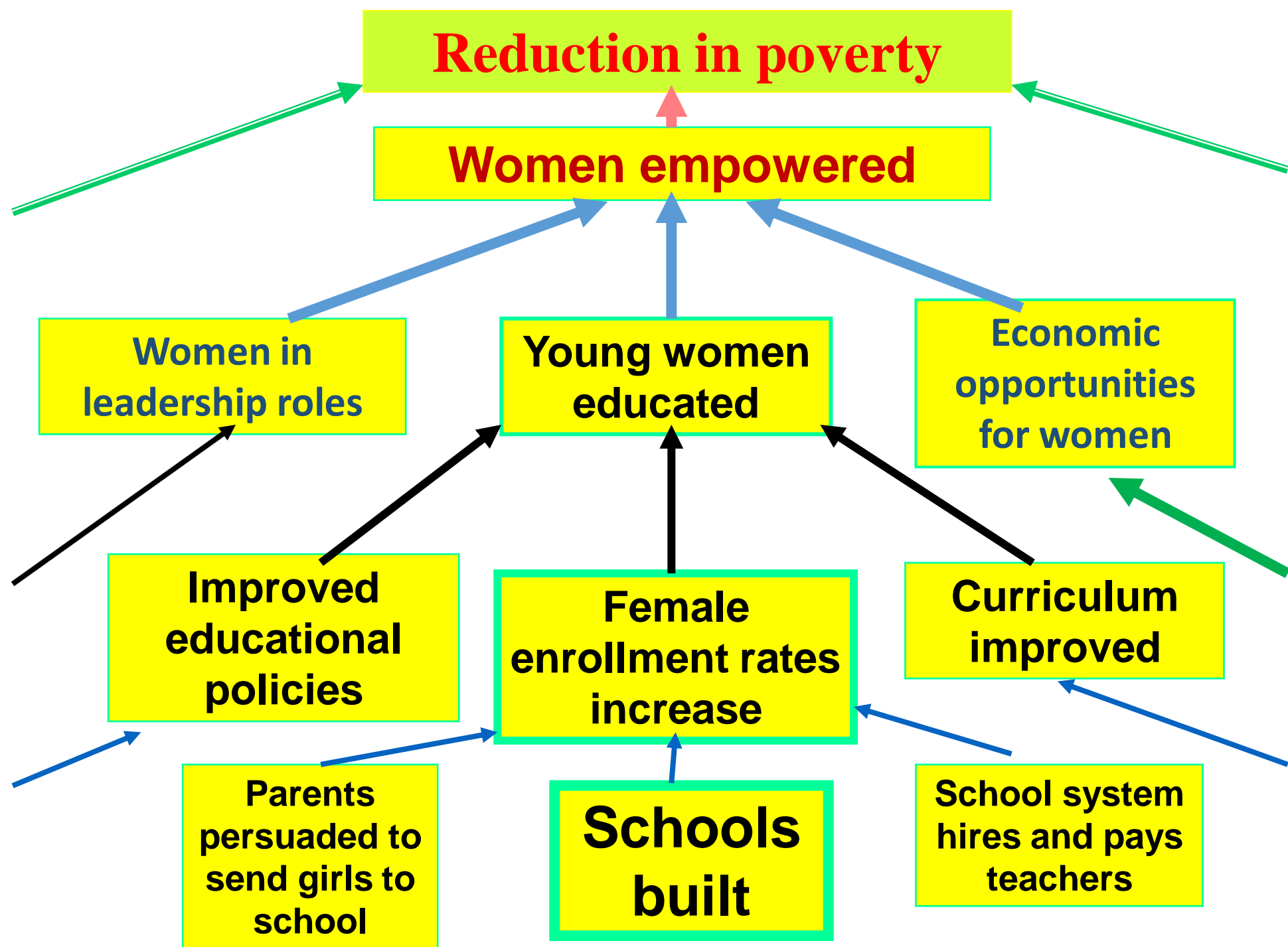
To have synergy and achieve impact all of these need to address the same target population.

Program Goal: Young women educated

Advocacy Project Goal: Improved educational policies enacted

Construction Project Goal: More classrooms built

Teacher Education Project Goal: Improve quality of curriculum

ASSUMPTION
(that others will do this)

OUR project

PARTNER will do this

*Program goal at **impact** level*
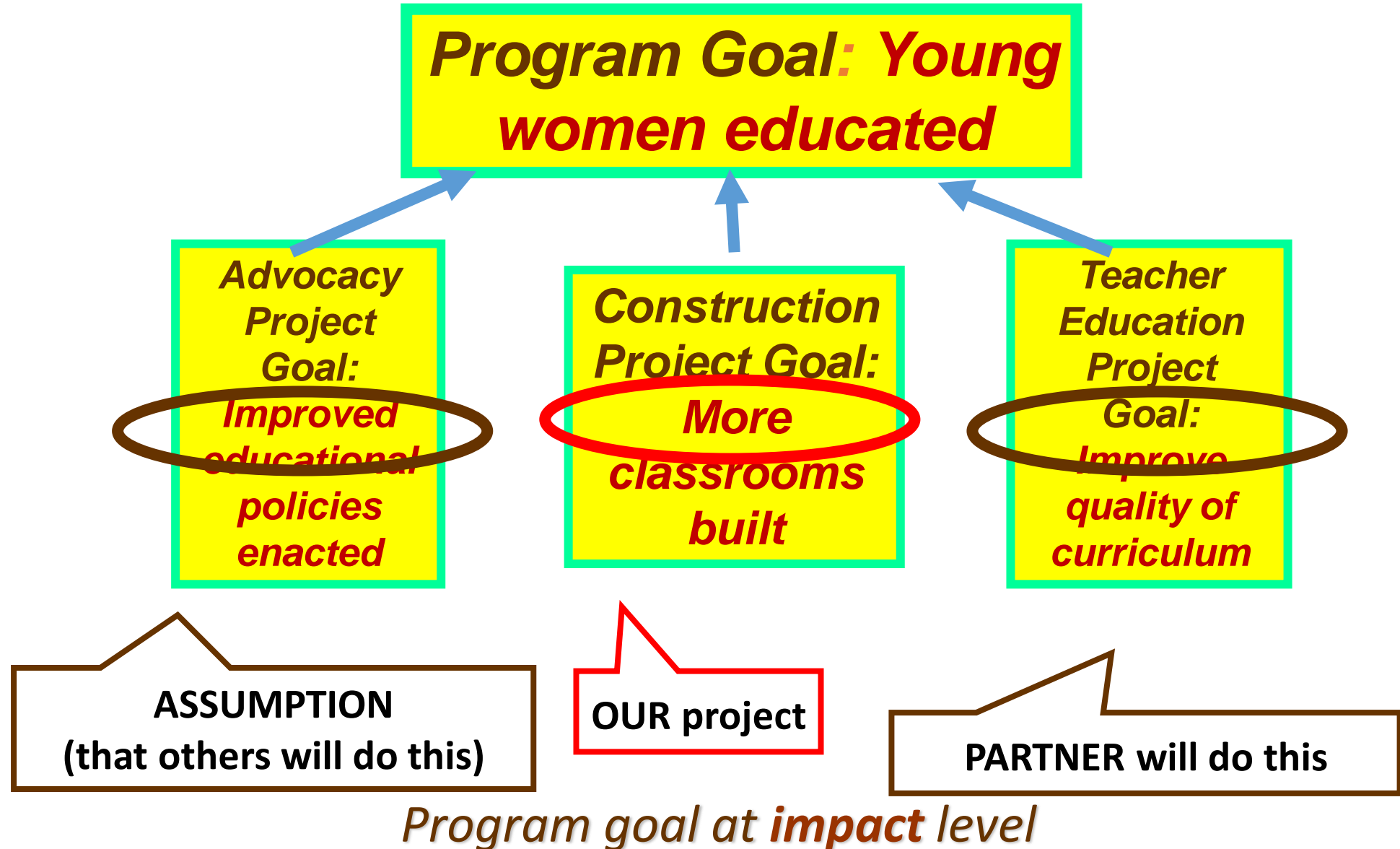
What does it take to measure indicators at each level?

Impact : Population-based survey (baseline, endline evaluation)

Outcome: Change in behavior **of participants** (can be surveyed annually)

Output: Measured and reported by project staff (annually)

Activities: On-going (monitoring of interventions)

Inputs: On-going (financial accounts)

We need to recognize <u>which</u> evaluative process is most appropriate for measurement at various levels

- **Impact**

- **Outcomes**

- **Output**

- **Activities**

- **Inputs**

IMPACT EVALUATION

PROJECT EVALUATION

MONITORING SYSTEM

FINANCIAL SYSTEM
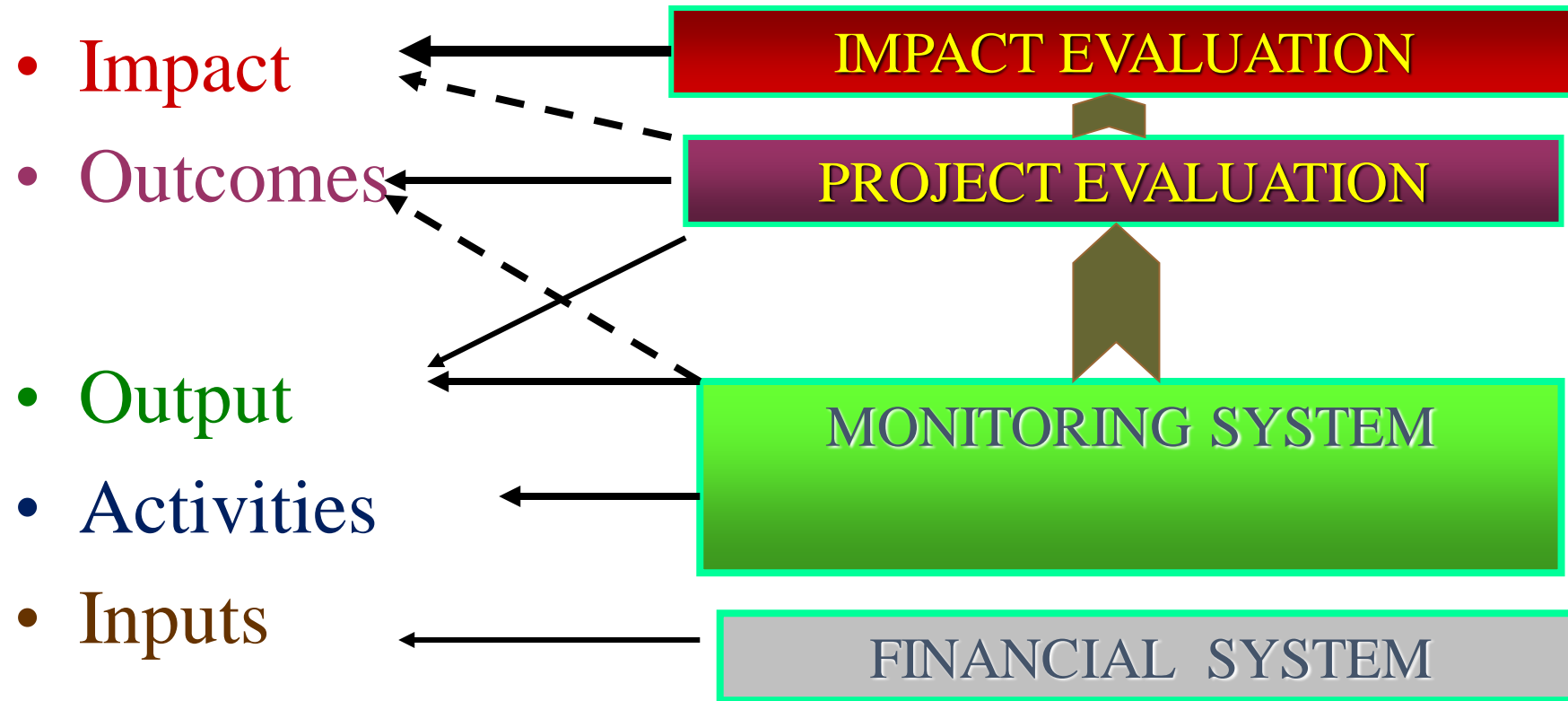
Another form of Program Theory (Logic) Model

| Economic context in which the project operates | Political context in which the project operates | Institutional and operational context |

| Design | Inputs | Implementation Process | Outputs | Outcomes | Impacts | Sustainability |

Socio-economic and cultural characteristics of the affected populations
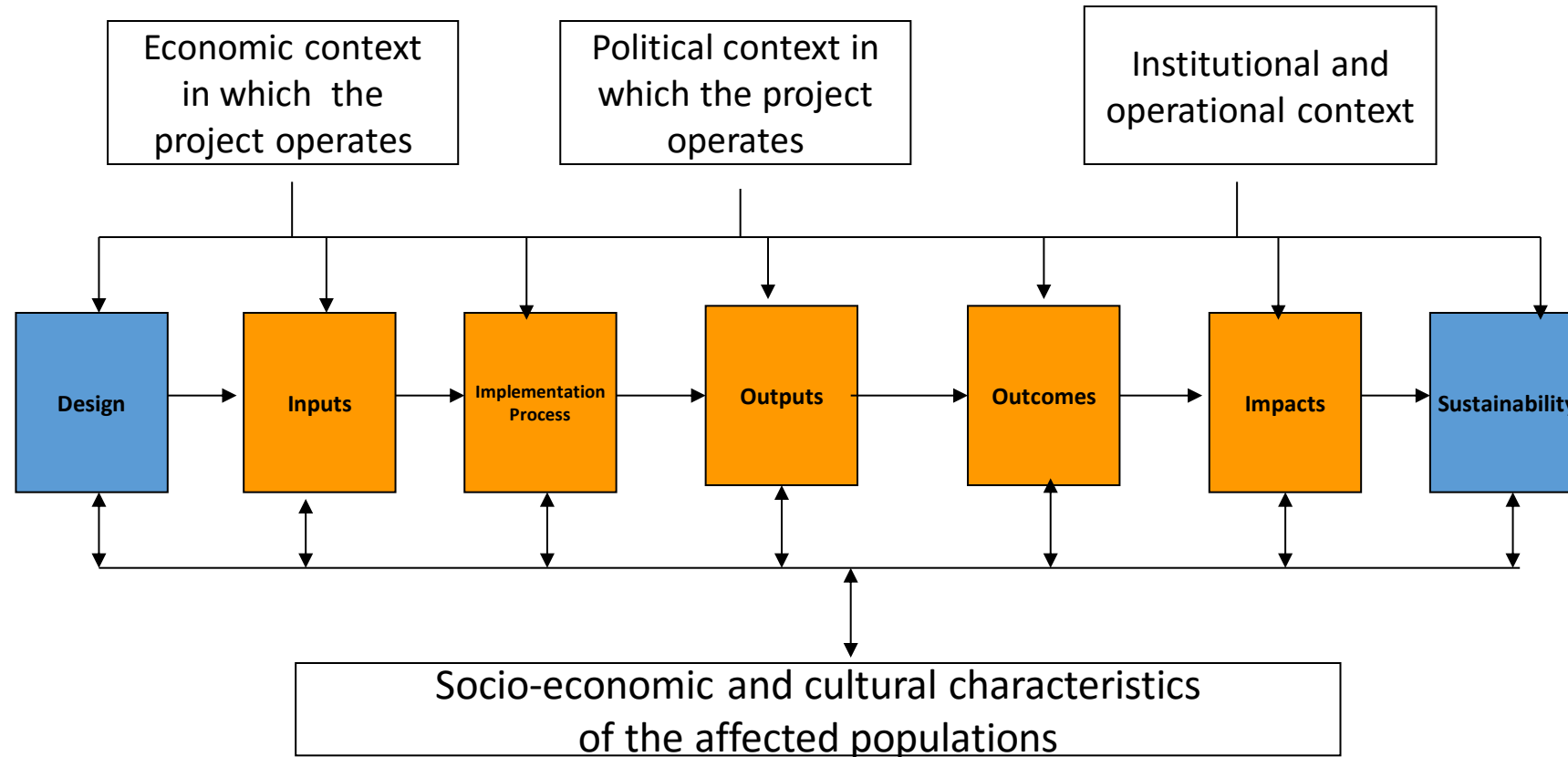
Note: The **orange** boxes are included in conventional Program Theory Models. The addition of the **blue** boxes provides the recommended more complete analysis.

Expanding the results chain for multi-donor, multi-component program

Impacts

Increased rural H/H income	Increased political participation	Improved education performance	Improved health

Intermediate outcomes

Increased production	Access to off-farm employment	Increased school enrolment	Increased use of health services

Outputs

Credit for small farmers	Rural roads	Schools	Health services

Inputs

Donor	Government	Other donors

Attribution gets very difficult! Consider ***plausible contributions*** each makes.

Sustainable Development Goals (SDGs)

Goal 1: **End poverty in all its forms everywhere**

Goal 2: **End hunger, achieve food security and improved nutrition and promote sustainable agriculture**

Goal 3: Ensure healthy lives and promote well-being for all at all ages

Goal 4: Ensure inclusive and quality education for all and promote lifelong learning

Goal 5: **Achieve gender equality and empower all women and girls**

Goal 6: **Ensure access to water and sanitation for all**

Goal 7: **Ensure access to affordable, reliable, sustainable and modern energy for all**

Goal 8: **Promote inclusive and sustainable economic growth, employment and decent work for all**

Goal 9: **Build resilient infrastructure, promote sustainable industrialization and foster innovation**

Sustainable Development Goals (SDGs)

Goal 10: Reduce inequality within and among countries

Goal 11: Make cities inclusive, safe, resilient and sustainable

Goal 12: **Ensure sustainable consumption and production patterns**

Goal 13: Take urgent action to combat climate change and its impacts

Goal 14: Conserve and sustainably use the oceans, seas and marine resources

Goal 15: Sustainably manage forests, combat desertification, halt and reverse land degradation, halt biodiversity loss
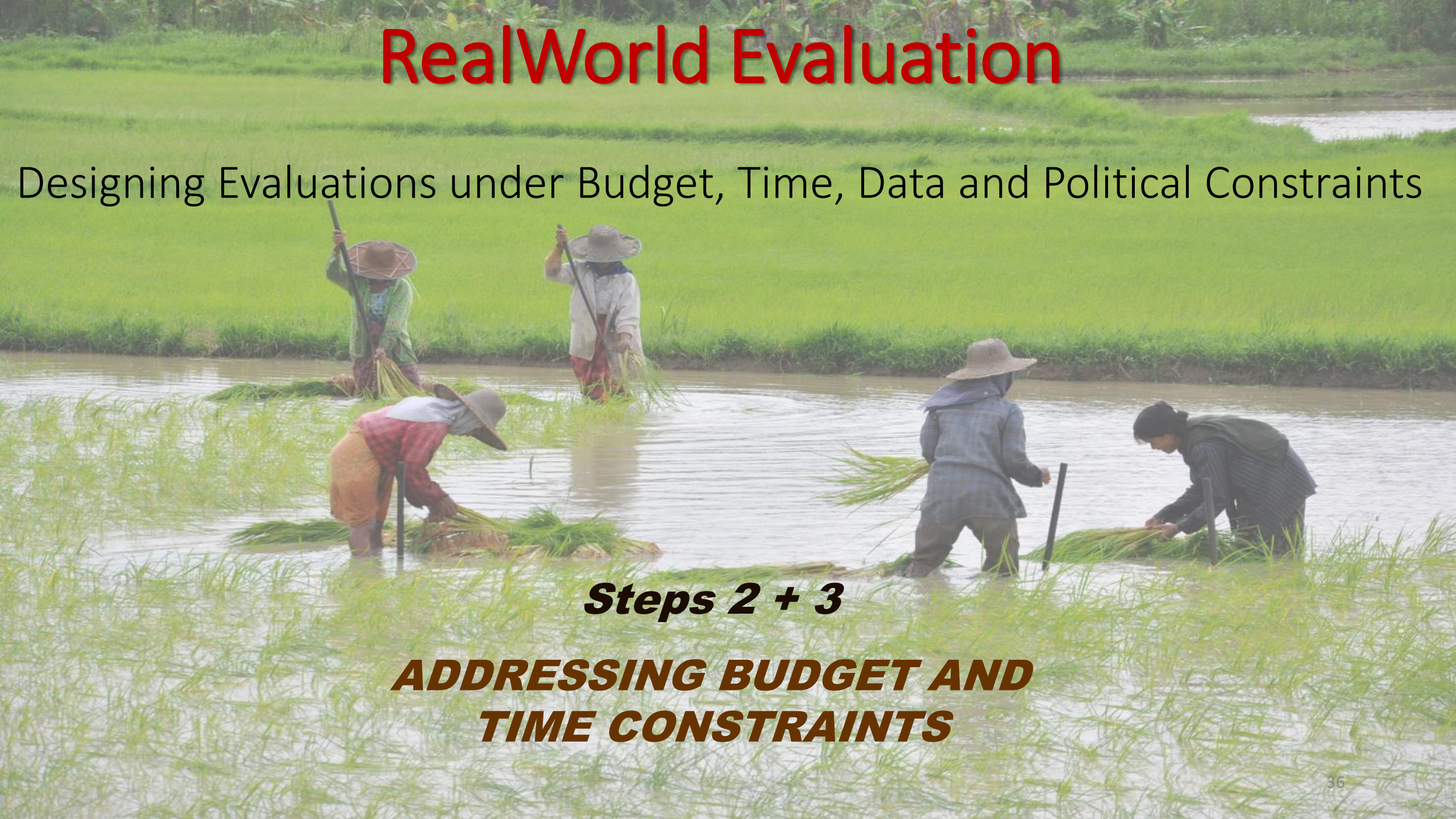
Goal 16: Promote just, peaceful and inclusive societies

Goal 17: Revitalize the global partnership for sustainable development

RealWorld Evaluation

Designing Evaluations under Budget, Time, Data and Political Constraints

Steps 2 + 3

ADDRESSING BUDGET AND TIME CONSTRAINTS

A. Clarifying information needs

- Who are the key stakeholders with interest in this evaluation, and what are they really expecting from this evaluation?

- What are the most important questions to be answered?

- How rigorous (detailed) does the evaluation need to be?

C . Look for reliable secondary sources

Assess the relevance and reliability of sources for the evaluation with respect to:

- Coverage of the target (or comparison) population
- Time period (when was it collected?)
- Relevance of the information collected
- Reliability and completeness of the data

D. Seeking ways to reduce sample size

Accepting a lower level of precision significantly reduces the required number of interviews:

- To test for a **5%** change in proportions (95% confidence level) requires a random sample of **1086**

- To test for a **10%** change in proportions (90% confidence level) requires a random sample of **270**

D. Seeking ways to reduce sample size, cont.

- Accept a lower level of statistical precision
- Reduce the number of levels of disaggregation of the analysis (different characteristics of target population)

Addressing time constraints

Negotiate with the client to discuss questions such as the following:

1. What information is essential and what could be dropped or reduced?

2. How much precision and detail is required for the essential information? E.g. is it necessary to have separate estimates for each geographical region or sub-group or is a population average acceptable?

3. Is it necessary to analyze all project components and services or only the most important?

4. Is it possible to obtain additional resources (money, staff, computer access, vehicles etc) to speed up the data collection and analysis process?

RealWorld Evaluation

Designing Evaluations under Budget, Time, Data and Political Constraints
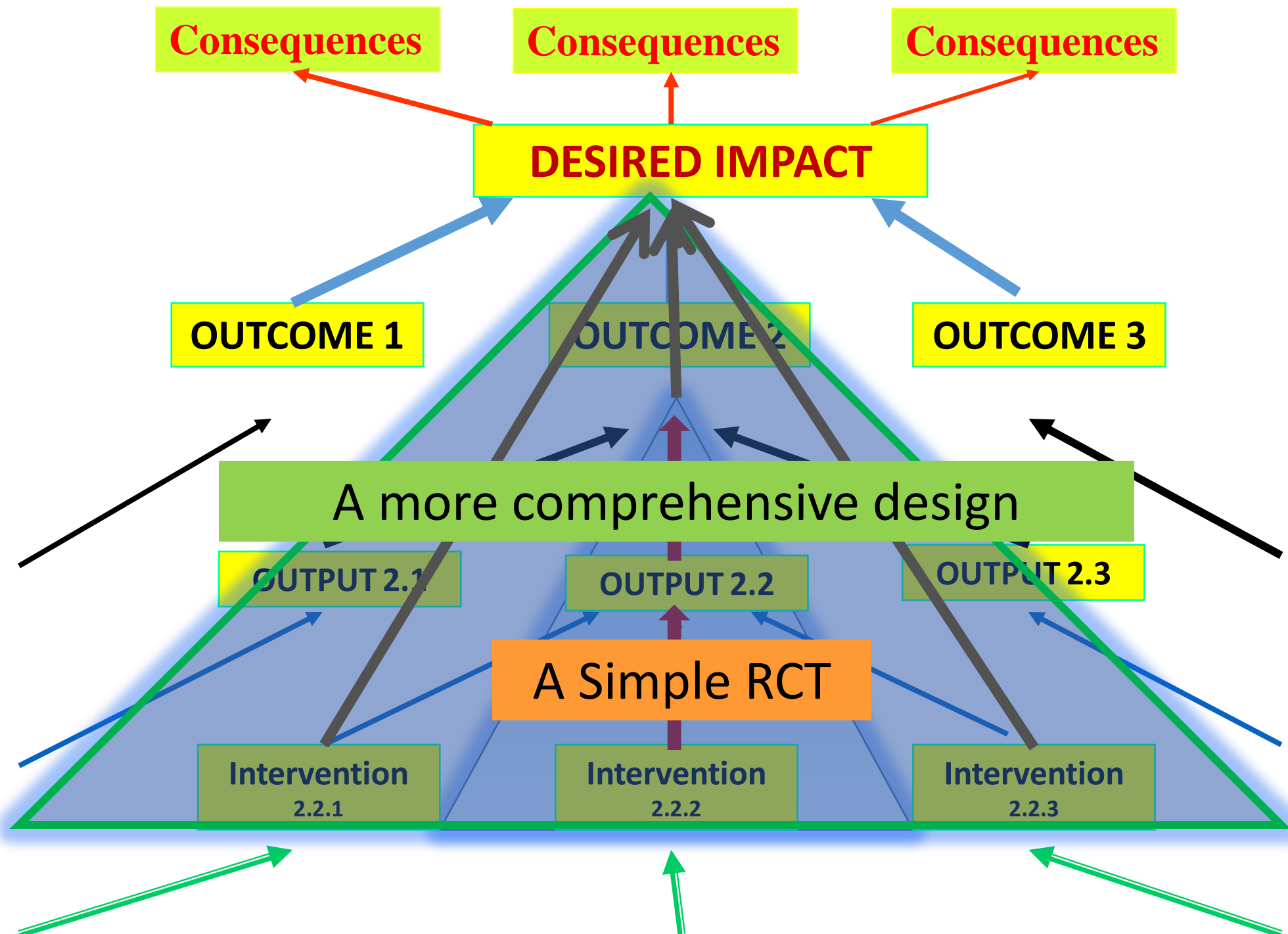
Determining Counterfactuals

There a variety of methods for assessing the counterfactual

- Collecting data from comparison groups (that were not reached by the project)
- Reliable secondary data that depicts relevant trends in the population
- Longitudinal monitoring data (if it includes non-reached population)
- Qualitative methods to obtain perspectives of key informants, participants, neighbors, etc.

Ways to reconstruct comparison groups

- Judgmental matching of communities
- When there is phased introduction of project services beneficiaries entering in later phases can be used as "pipeline" comparison groups
- Internal controls when different subjects receive different combinations and levels of services

Consequences · Consequences · Consequences

DESIRED IMPACT

OUTCOME 1 · OUTCOME 2 · OUTCOME 3

A more comprehensive design

OUTPUT 2.1 · OUTPUT 2.2 · OUTPUT 2.3

A Simple RCT

Intervention 2.2.1 · Intervention 2.2.2 · Intervention 2.2.3

There are two main questions to be answered by "impact evaluations":

1. Is the quality of life of our intended beneficiaries improving?
2. Are our programs making *plausible contributions* (along with other influences) towards such positive changes?

The 1st question is much more important than the 2nd!

Thank you!

Main workshop messages

1. **Evaluators must be prepared for RealWorld evaluation challenges.**

2. **There is considerable experience to learn from.**

3. **A toolkit of practical "RealWorld" evaluation techniques is available (see www.RealWorldEvaluation.org).**

4. **Never use time and budget constraints as an excuse for sloppy evaluation methodology.**

5. **Evaluators and clients need to be clear on main purpose for evaluation, and then what design and methods would be feasible and adequate.**