



USAID | **ETHIOPIA**
FROM THE AMERICAN PEOPLE

Lesson 8: Introduction to Evaluation



Learning Objectives

- 1) Provide an overview of evaluation
- 2) Understand USAID Policies and Practices
- 3) Understand key steps in planning and conducting evaluations
- 4) Know the basic principles behind developing a good statement of work



Global Trends in Development Evaluation

- Aid effectiveness Forum: Paris Declaration (2005), Accra (2008), Busan (2011)
- Collaborative evaluations and harmonization on results.
- Impact evaluation to inform policy making.
- Evaluation capacity development as good governance.
- Higher plan, thematic and sector evaluations.
- Transparency and evaluation as a public good.





USAID | **ETHIOPIA**
FROM THE AMERICAN PEOPLE

Evaluation: USAID Context

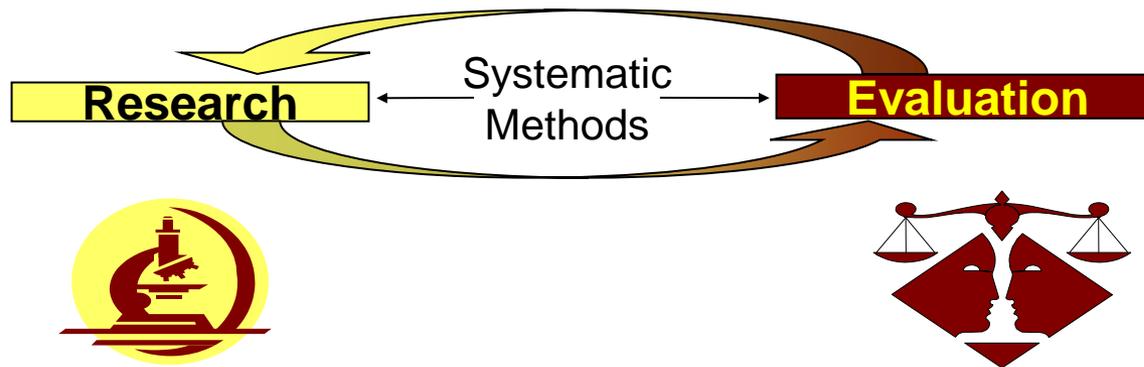


What is Evaluation?

- **Evaluation** is the *systematic collection and analysis* of information about the *characteristics and outcomes* of programs and projects as *a basis for judgments*, to *improve effectiveness*, and/or *inform decisions* about current and future programming. (ADS 200).
- Evaluation is distinct from assessment, which may be designed to examine country or sector context to inform project design, or an informal review of projects. (Chapters 200-203).



Research vs. Evaluation



- Production of generalizable knowledge
- Researcher-derived questions
- Paradigm stance
- More controlled setting
- Clearer role
- Published
- Clearer allegiance

- Knowledge intended for use
- Program- or funder-derived questions
- Judgmental quality
- Action setting
- Role conflicts
- Often not published
- Multiple allegiances



“Research seeks to *prove*,
evaluation seeks to *improve*...”

M. Q. Patton

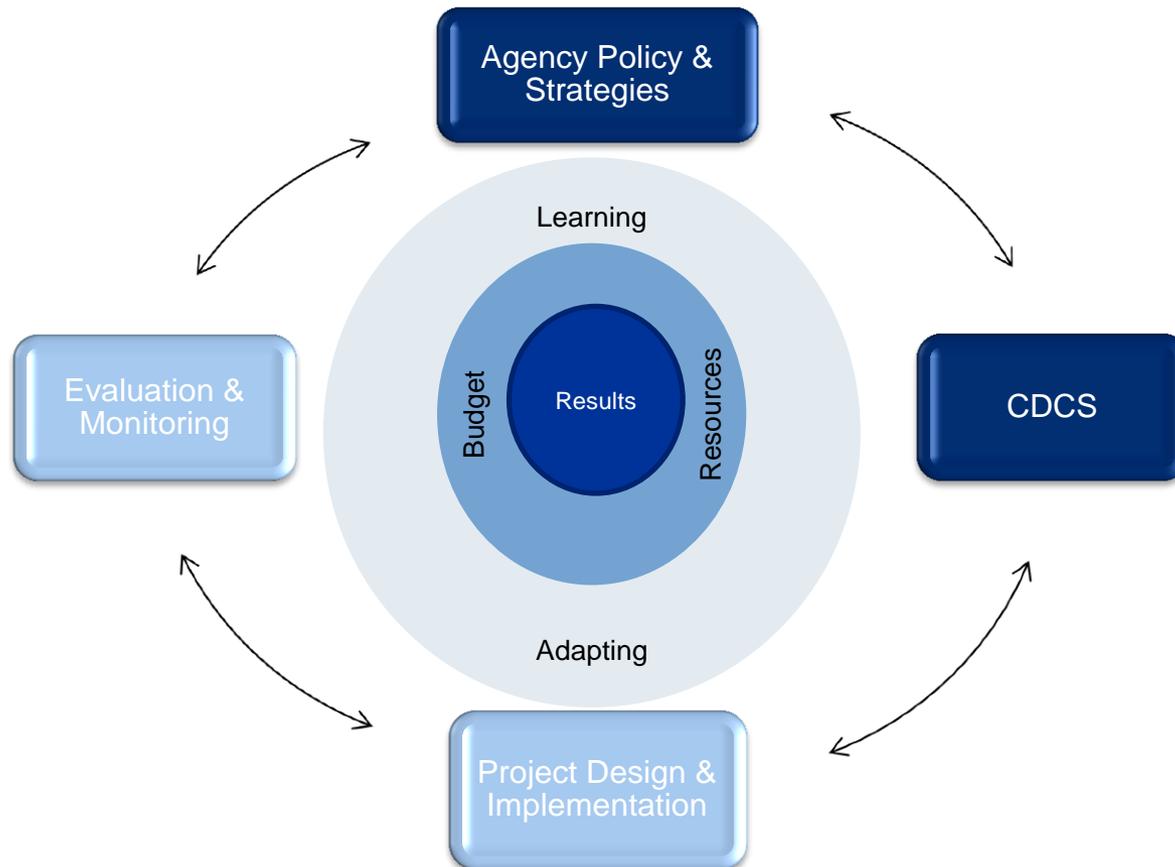


Why Evaluation?

- **Evaluation in USAID has two primary purposes:** accountability to stakeholders and learning to improve development program/project effectiveness.
 - To **gain insight** about a program and its operations – to see where we are going and where we are coming from, and to find out what works and what doesn't
 - To **improve practice** – to modify or adapt practice to enhance the success of activities
 - To **assess effects** – to see how well we are meeting objectives and goals, how the program benefits the community, and to provide evidence of effectiveness
 - To **build capacity** - increase funding, enhance skills, strengthen accountability



USAID Program Cycle





Current Policies and Practices

Integrated into Project Design

- For CDCS countries identify opportunity for impact evaluation for each DO.
- Identify key evaluation questions at outset.
- Considerations such as providing budget for evaluation and baseline data needs.
- Include evaluation specialists in design teams, and plan for baselines

Unbiased

- Implementing partners do not evaluate own projects. (See Eval Policy FAQ)
- USAID staff can serve on an evaluation team, but must use external (to USAID) team lead for all evaluation teams.
- Disclose Conflicts of Interest

Relevant

- Link evaluation questions to future decisions.
- Consult with in-country partners and beneficiaries.
- Use findings for program management and project design.



Current Policies and Practices

Based on Best Methods

- Use methods that generate replicable and high quality evidence given time, budget and other resource constraints.
- Qualitative and/or quantitative methods defined in advance in SOW.

Reinforcement of Local Capacity

- Include evaluation specialists from partner countries in evaluation teams and use host country systems where appropriate
- Place priority on building local capacity to undertake evaluations and use the results generated.

Transparency

- Findings and data will be shared widely.
- Make evaluation findings available to the public through USAID website within three months of completing final draft.

C-1: Checklist for Defining “Quality Evaluation”



When Does USAID Require Evaluation ?

Required

Large projects: at or above average dollar value for projects within each development objective managed by an OU

Innovative or pilot projects of any size: demonstrate new approaches that are anticipated to be expanded in scale or scope if the approach is proven successful.

Recommended

(in addition to large and pilot projects)

Evaluations at the **program or sector level**, particularly valuable in a period preceding the development of a new strategy.

Any other evaluations identified by an operating unit as needed for **learning or management purposes**.



Evaluation Decisions During Design

- **When a Mission develops its CDCS:**

“Operating units are encouraged to identify at least one opportunity for an impact evaluation for each DO during CDCS planning.”

- Evaluation for learning: Acknowledge what we don't know
- Impact Evaluation (IE) is not for the full DO, but something within it
 - ex/ DO: Increased agricultural productivity
 - IE: *What is the best way to encourage farmers to adopt enhanced fertilizer technology?*

Actual feasibility will be determined during design stage of the project.

Allocate 3% of program funds to evaluations!



Types of USAID Evaluations

Impact evaluations

- Measure the change in a development outcome that is attributable to a defined intervention. Impact evaluations are **based on models of cause and effect** and require a credible and rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change.

Performance evaluations

- Incorporate before-after comparisons, but generally lack a rigorously defined counterfactual. Performance evaluations **focus on descriptive and normative questions**: what a particular project or program has achieved (either at an intermediate point or at the conclusion of an implementation period); how it is being implemented; how it is perceived and valued; whether expected results are occurring; and other questions that are pertinent to program design, management and operational decision making.



USAID | **ETHIOPIA**
FROM THE AMERICAN PEOPLE

Deciding to Evaluate Program/Project



Evaluation Decisions During Design

•During Project Design

- Plan for baseline
- Determine feasibility of IE if planned

In a Performance Management Plan (PMP):

- Include possible evaluation efforts to complement performance monitoring
- Estimate the costs of collecting, analyzing, and reporting performance data, and plan how these will be financed.

When developing a project cost estimate and financial plan, USAID teams must include the costs of **evaluation**.



USAID Evaluation “Triggers” During Implementation

1

- A key management decision is required, but there is inadequate information to make it.

2

- Performance information indicates an unexpected result (positive or negative) that should be explained, such as unanticipated results affecting either men or women.

3

- Customer, partner, or other informed feedback suggests that there are implementation problems, unmet needs, or unintended consequences or impacts.

4

- Issues of sustainability, cost-effectiveness, or relevance arise.

5

- The validity of Results Framework hypotheses or critical assumptions is questioned; for example, due to unanticipated changes in the host country environment.

6

- Periodic Portfolio Reviews have identified key questions that need to be answered or that require consensus.
- Extracting lessons is important for the benefit of other ongoing or future programs, here or elsewhere.

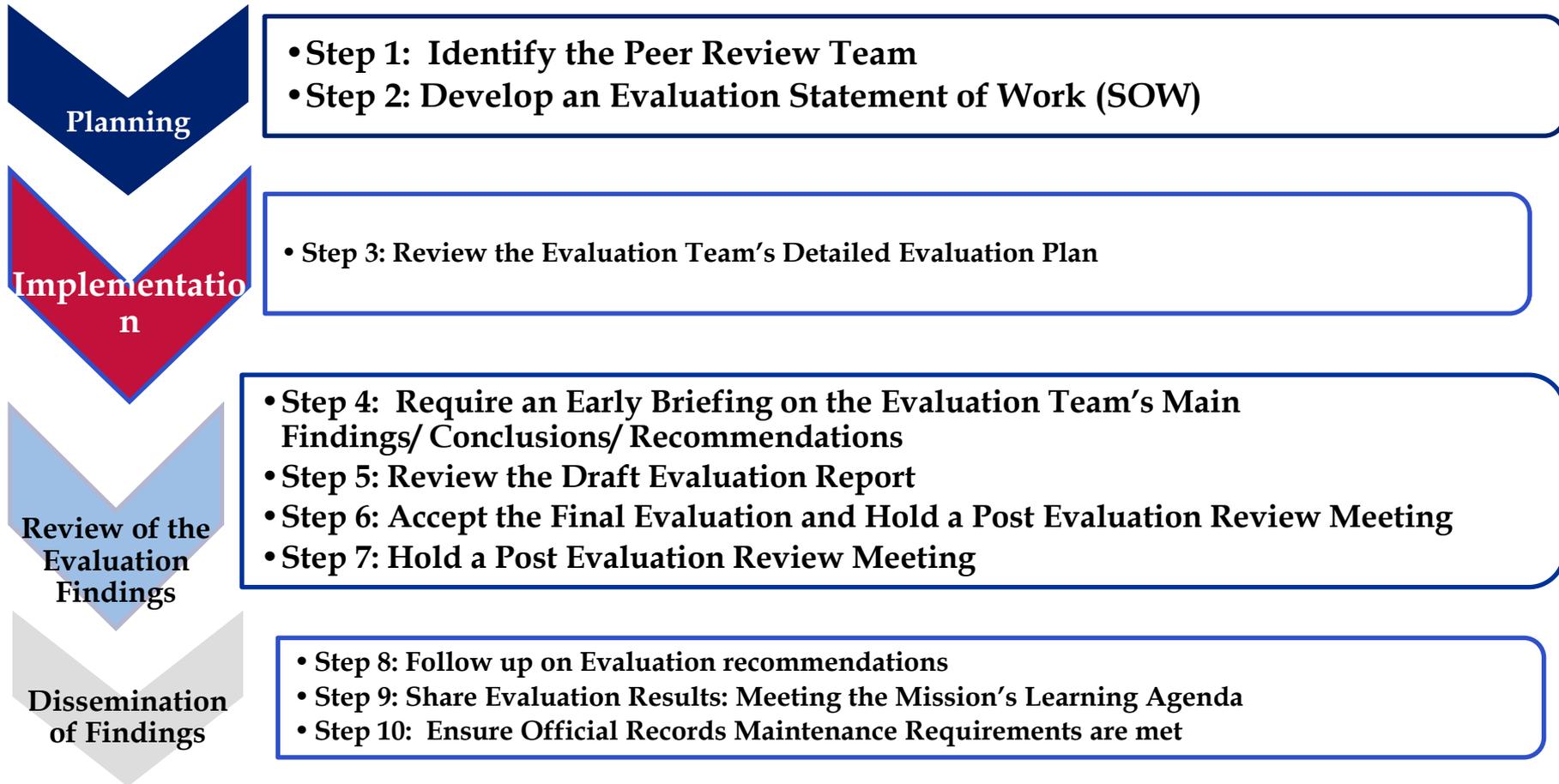


USAID | **ETHIOPIA**
FROM THE AMERICAN PEOPLE

Steps in Program Evaluation Process



Managing the Evaluation Process





Underlying Logic of Evaluation Process

- ***No evaluation is good unless...*** findings are ***used*** to make a difference
- ***No findings are used unless...*** a ***demand*** has been created prior to creating the product
- ***No demand is created unless...*** the evaluation is ***well-focused***, including most relevant and useful questions



The Four Standards

- **Utility:** Who needs the information and what information do they need?
- **Feasibility:** How much money, time, and effort can we put into this?
- **Propriety:** What steps need to be taken for the evaluation to be ethical?
- **Accuracy:** What design will lead to accurate information?



USAID | **ETHIOPIA**
FROM THE AMERICAN PEOPLE

Designing an Evaluation: Key Considerations and Components

Evaluation Purpose



Purpose – Intended Users

The primary intended users of the evaluation

- USAID decision-makers
- The project implementation team

Important secondary audiences for the evaluation

- The government of the assisted country
- USAID staff in other countries who might want to design a similar project



USAID | **ETHIOPIA**
FROM THE AMERICAN PEOPLE

Designing an Evaluation: Key Considerations and Components

Defining Evaluation Questions



Good Evaluation Questions

Mid Term	Final Evaluations	Impact & Sustainability
How the startup phase was operationalized?	Whether the benefits people received were the same as those the activity planned to deliver?	Is there any evaluative evidence (qualitative or quantitative information) regarding the impact of the activity?
Whether implementation processes for solving problems and/or delivery of inputs are in place and functioning efficiently	What factors were most important for ensuring that people received benefits from the activity?	Can the causal links between the intervention and their effect be explained?
Whether initial expectations about how the activity would interact with its intended beneficiaries are being fulfilled	What impediments had to be overcome, or were never overcome and why?	What were the lessons learned, both positive and negative?
	Were there any unplanned results, whether positive or negative?	Did the activity produce any spin-off benefits?
	Were the activity costs as planned or did they exceed or were they lower than expected and why?	Did other communities use it as a model for their own projects?
	Is there any evidence that suggests that the activity would produce these same results if it were tried elsewhere? This deals with issues of replicability.	



Identifying and Selecting Questions

Two phases:

- *Divergent Phase:*
develop a comprehensive list of questions
- *Convergent Phase:*
narrow down the list





Sub-questions are Often Needed

How relevant was the intervention?

To what extent did the project fit within USAID's country health strategy?

To what extent was the intervention aligned with the host country government's priorities?

Was the timing appropriate given the political context?

What were the characteristics of those served?

What were the most prevalent age groups?

To what extent was gender equity achieved?

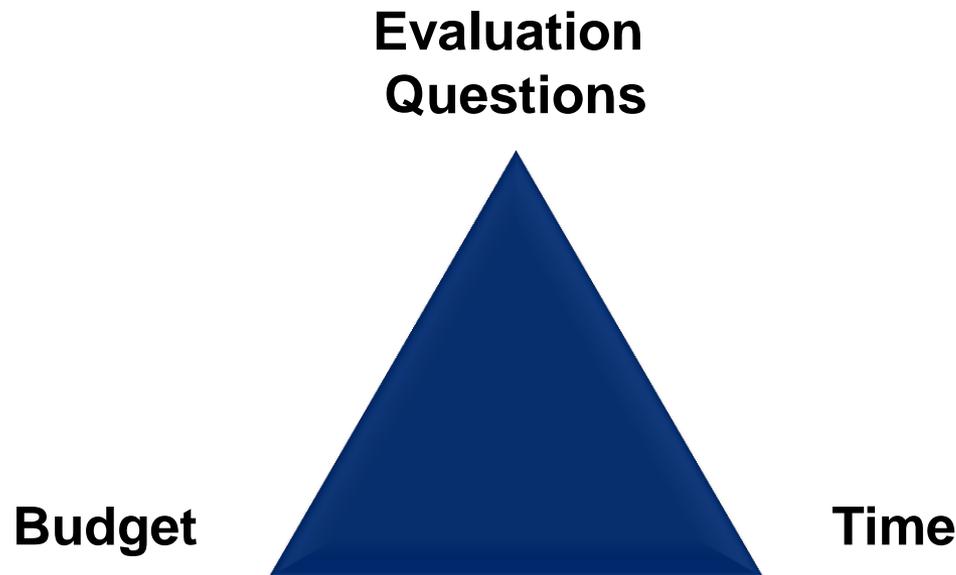
What were participants' most frequent health issues?

Did we target the youth most considered 'at risk'?



Balance

A sound balance between evaluation questions and the time and resources available to answer them is essential.





A Final Point

Defining evaluation questions should be a collaborative and iterative process that serves to promote later use of evaluation findings!





USAID | **ETHIOPIA**
FROM THE AMERICAN PEOPLE

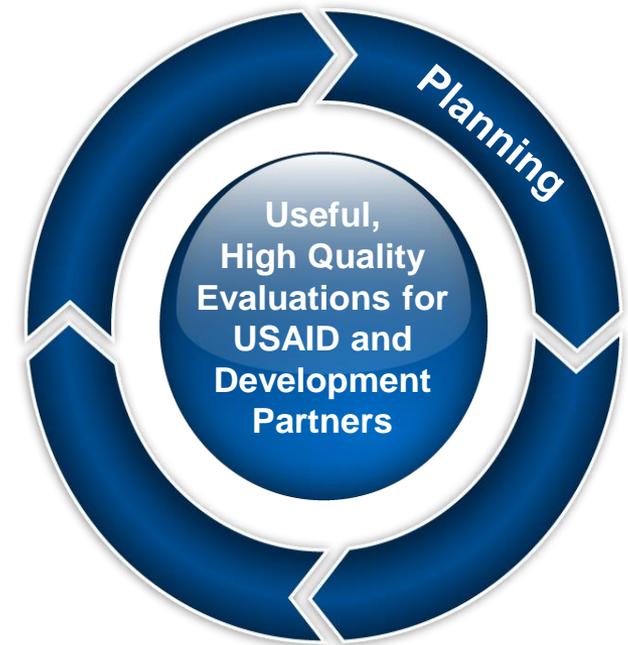
Designing an Evaluation: Key Considerations and Components

Choosing an Evaluation Approach/Design



The Right Design is Key to Planning a Successful Evaluation

- When is an evaluation design the “right design”?
- When it is directly linked to the evaluation questions
- When its selection is based on sound information regarding the what’s, why’s and how’s of the intervention to evaluate
- When it takes validity into account





Evaluation Design Continuum

Focus on causality



Focus on implementation issues

Performance Evaluation Designs

Impact Evaluation Designs

Non experimental designs

Weaker QEDs

Stronger QEDs

Experimental designs

Designs without comparison groups or randomized assignment

Example:
Performance Evaluations that focus on implementation processes and efficiency (descriptive and normative or comparative questions)

Designs with comparison groups – but not randomized assignment

Example:
•Performance Evaluation that marginally focus on results, by answering questions such as “Did the program achieve its objectives?”

The strength of this design depends on the technical rigor of the technique used to identify the comparison group

Designs with randomized assignment (inclusion of a control group) to definitively answer cause-effect questions

QED = quasi-experimental design





The Timing of Different Evaluations

Evaluation Type	Timing
Impact Evaluation	<p>Produced at or after project completion and used to plan or scale-up future projects.</p> <ul style="list-style-type: none">• IEs should be designed during the earliest stages of program design.• Some IEs can be designed later in the program though many more technical challenges are involved.
Performance Evaluation	<p>A range of evaluation options that can be produced at any phase of the program cycle to improve performance.</p> <ul style="list-style-type: none">• All projects will consider the design of PEs during the project design stage. This is part of the preparation of a performance management plan.• All PEs should be commissioned well in advance of the time frame for program/strategy decisions.



When to design a Performance Evaluation?

- At any phase of the program cycle as long as it will help to improve the performance of the project/program to evaluate
- Often during the project design stage (e.g., as part of the Performance Management Plan)
- In any case, well in advance of the program/strategy decisions informed by the evaluation findings





USAID | **ETHIOPIA**
FROM THE AMERICAN PEOPLE

Designs for Impact Evaluation



USAID Definition of Impact Evaluation

- *Impact evaluations measure the change in a development outcome that is attributable to a defined intervention; **impact evaluations are based on models of cause and effect and require a credible and rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change.** Impact evaluations in which comparisons are made between beneficiaries that are randomly assigned to either a “treatment” or a “control group provide the strongest evidence of a relationship between the intervention under study and the outcome measured.*



USAID Definition of Impact Evaluation

- *Impact evaluations measure the change in a development outcome...*
 - Focused on higher level results (outcomes) rather than inputs and outputs
 - E.g. health, income, school enrolment, agricultural productivity
 - Over a given period of time to measure change
- *...that is attributable to a defined intervention*
 - What do we mean by intervention?
 - The level is flexible...activity, project, or program
 - BUT, it must be clearly defined and implemented
 - Must rule out other potential causes

But how do we do that?

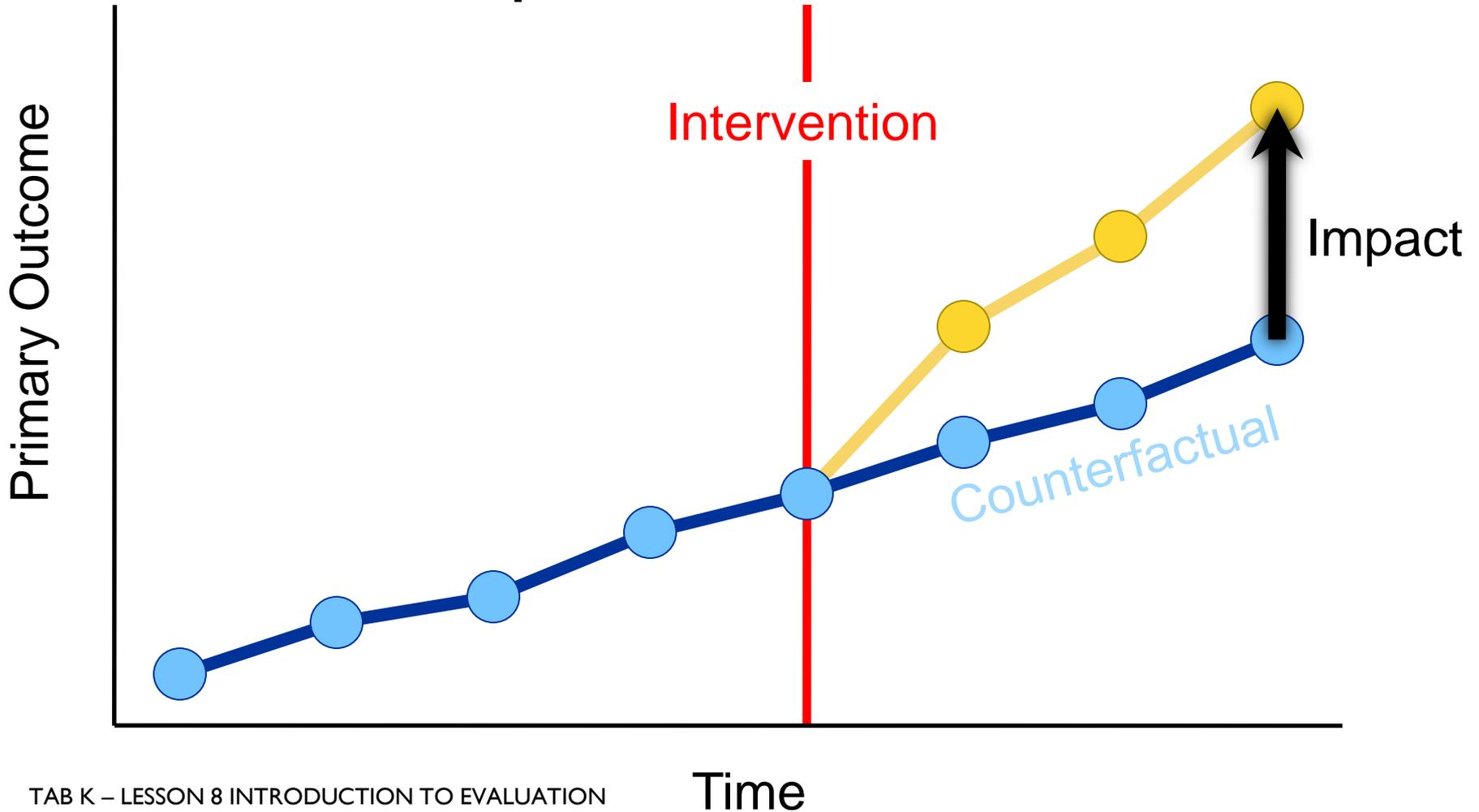


USAID Definition of Impact Evaluation

- *Impact Evaluations are based on models of cause and effect ...*
 - Requires a causal theory of change
 - E.g. results framework, logical framework, development hypothesis
- *... and require a credible and rigorously defined **counterfactual** to control for factors other than the intervention that might account for the observed change.*
 - The **counterfactual**... *identifies what would have happened to the beneficiaries absent the program.*
 - The 'impact' of a program is defined as changes relative to this counterfactual



Counterfactual...in pictures





Change Does Not Prove Causality

An observed
change

≠

Causality
[impact]

More children
attending
school in project
areas

Does **NOT**
prove

Project interventions
caused the increase



Main Types of Impact Evaluation Designs

1. Experimental Designs (Randomized Control Trials):

- Units are randomly assigned to the project and control groups
- This is the strongest statistical design as it provides unbiased estimates of project impacts

2. Quasi-Experimental Designs:

- Units are either self-selected or selected by the project agency
- A comparison group is selected to match as closely as possible the project group
- Design quality varies in terms of how closely the two groups are matched (**selection bias**)

3. Non-Experimental Designs:

- Used when it is not possible to select a comparison group
- The analysis is statistically much weaker (subject to more sources of bias) but the findings may be credible and useful to managers



Impact Evaluation Questions

Which of the following are typical IE questions?

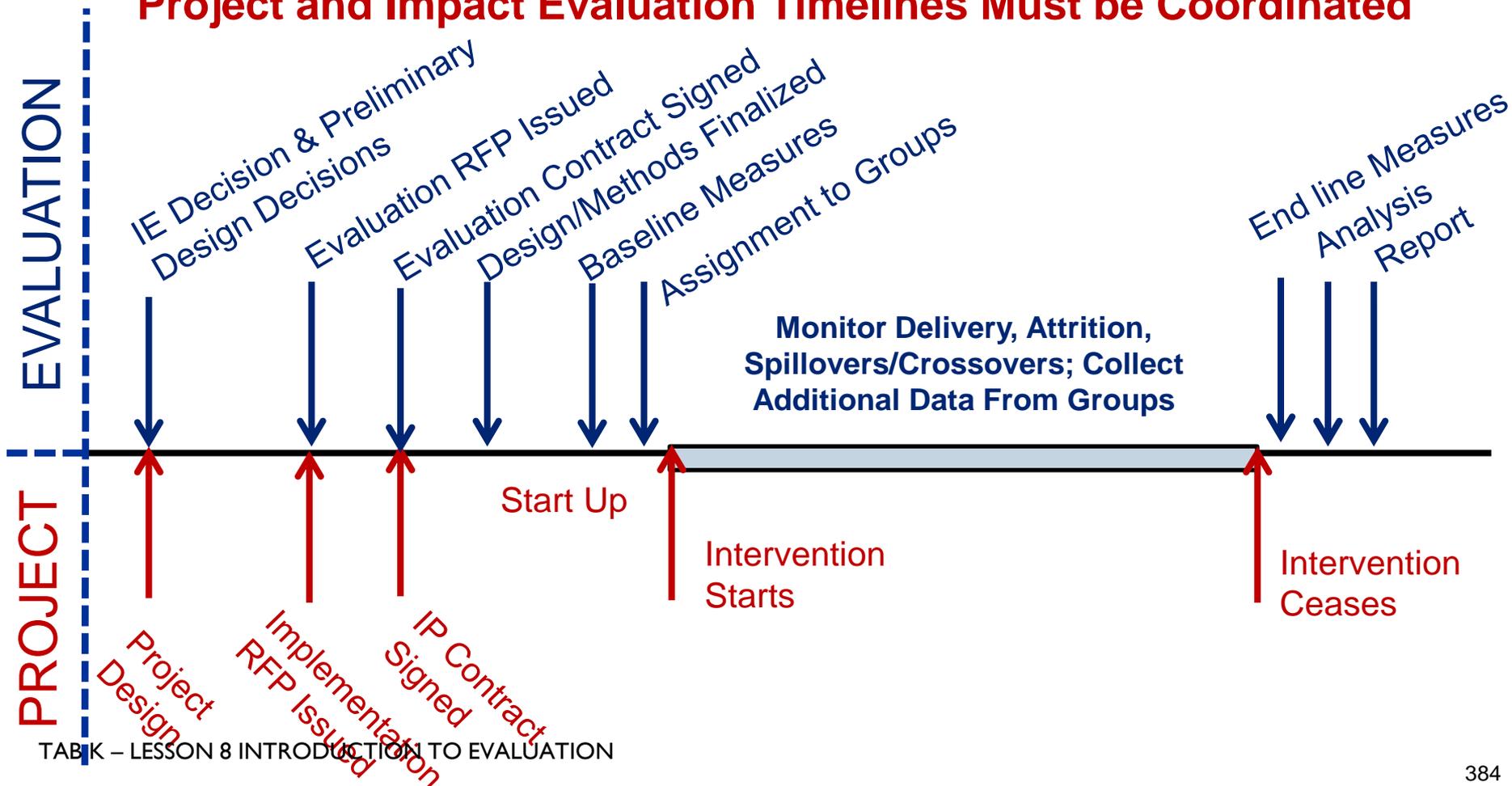
- Was it the intervention that caused observed results?
- What proportion of women participated in the program?
- Were women affected differently than men?
- Was the program implemented on time and budget?
- Which approach is most effective in achieving the desired result?
- Under what conditions would the intervention produce similar results if it were replicated?

Are these other questions still important for IEs?



Illustrative Prospective Impact Evaluation Timeline

Project and Impact Evaluation Timelines Must be Coordinated





USAID | **ETHIOPIA**
FROM THE AMERICAN PEOPLE

Designing an Evaluation: Key Considerations and Components

Developing an Evaluation SOW



Developing a SOW

A Statement of Work (SOW) is a defining document for an evaluation:



It should include all of the key substantive elements of what will become the agreement or contract between the sponsor of an evaluation and the evaluation team.



Developing a SOW

An Evaluation SOW provides structure – even for an internal evaluation team:

- It reminds an evaluation team of the important questions they need to answer;
- It helps a team organize its report, and
- It helps everyone involved stay focused on plans for utilizing the results of an evaluation.



Developing a SOW

It is important to understand precisely what you want to know from the evaluation and how you will get there before focusing on the elements of an SOW.

- What resources can be mobilized for the evaluation?
- Is there a decision timeline into which an evaluation must fit to be useful?
- Are certain questions and approaches ruled out by timing or resource constraints?
(For example, if it is Year 2 of the project, it may be too late to establish a valid comparison group unless the project is being phased-in to new districts each project year).



REVIEW: Organizing Questions for Writing a SOW

- What is the main purpose of this evaluation? (Accountability, Learning, or Other?)
- What specific information is needed to make a management decision?
- Who needs the information and why do they need it?
- What questions must be asked to obtain the required information?
- What are the data requirements? (Based on level of change expected from the intervention and level of certainty or precision required to make the decision?)
- What data is already available?
- What is the best design given information required and constraints? (Budget, timing, etc.)



Utilization is a Goal – Build Toward It

Treat the SOW development process as a first step towards the utilization of an evaluation.

- Consult with key stakeholders early to identify their interests and possible information requirements;
- Build awareness of the planned evaluation in USAID and among partners;
- Elicit input on the evaluation purpose, questions, and decision-making schedules from USAID staff and partners.



Elements of an Evaluation SOW

The Four Main Elements of a good evaluation SOW:

- **Program/Project Information** – what, where, when
- **Evaluation Fundamentals** – purpose, questions
- **Technical Requirements** – design, methods, staffing/competencies
- **Management Information** – schedule, budget, deliverables, POC/roles, logistics



Any Questions?

