Lecture III: Evaluation Design

Stanislao Maldonado

Universidad del Rosario
stanislao.maldonado@urosario.edu.co

Impact Evaluation
Universidad del Rosario
February 7th, 2017

1. Motivation

- Impact evaluation is closely related to program design
- Steps for preparing an impact evaluation (Gertler et al 2016):
  - Constructing a theory of change
  - Specifying the evaluation questions
  - Selecting indicators
- Before designing the evaluation, some previous steps are required to understand the logic of the intervention:

Front-end analysis

Investigation of an issue or problem to determine what is known about it and how to proceed in developing an evaluative approach to it

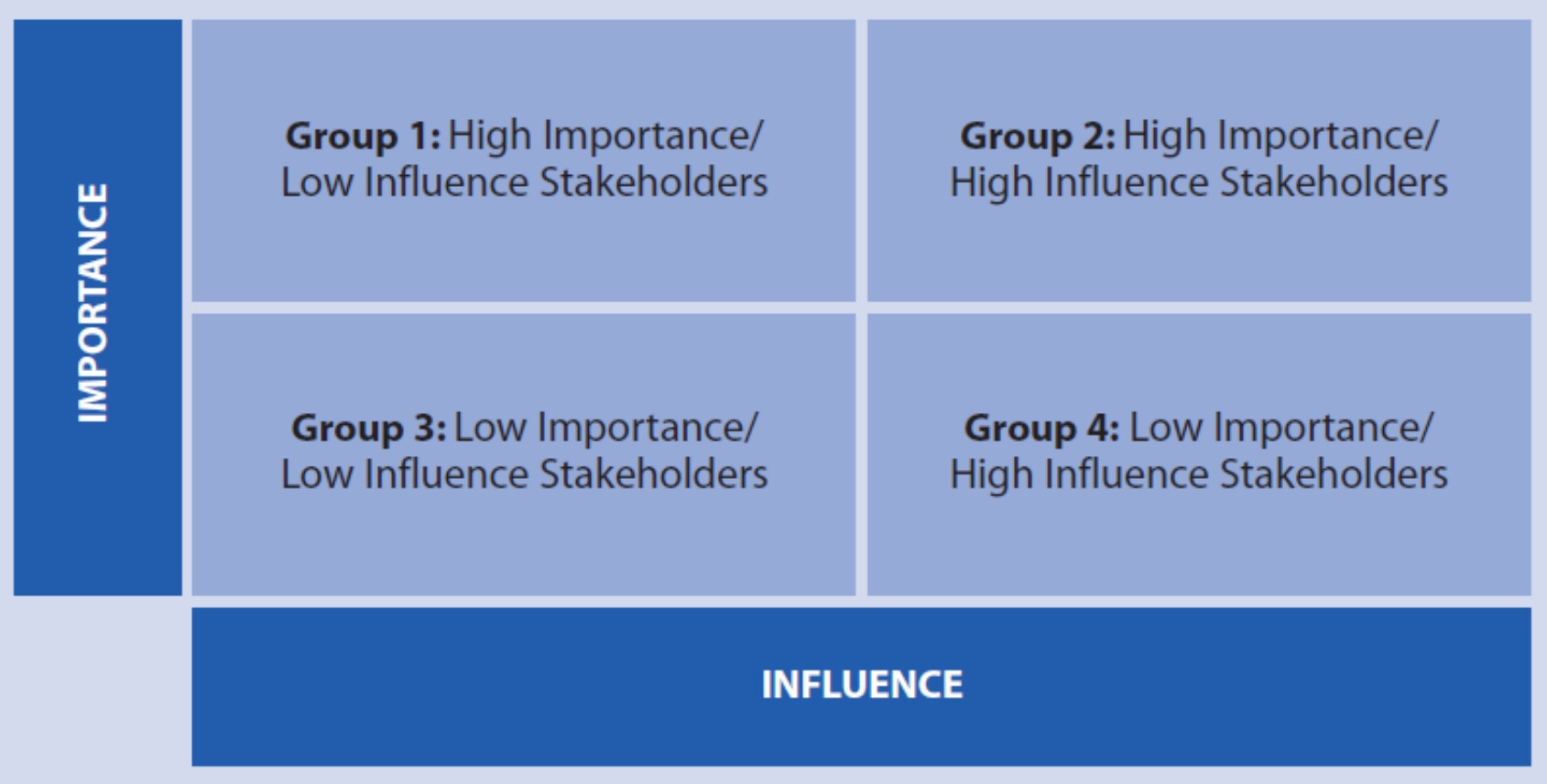
2. Front-End Analysis

- Who is the main client for the evaluation? Who are other important stakeholders? What issues do they identify for the evaluation?
- How will the timing of the evaluation in relation to the program implementation affect the evaluation?
- How much time is available for the evaluation?
- What is the nature and extent of available resources?
- Does social science theory have relevance for the evaluation?
- What have evaluations of similar programs found?
- What is the theory of change of the intervention?
- What existing data can be used for the evaluation?

Stakeholder analysis

- Identify people, group, and institutions that can influence the evaluation
- Anticipate the kind of influence these groups will have on the evaluation
- Develop strategies to get the most effective support possible for the initiative and to reduce obstacles to successful implementation of the evaluation

Stakeholder importance and influence matrix



3. Program Design

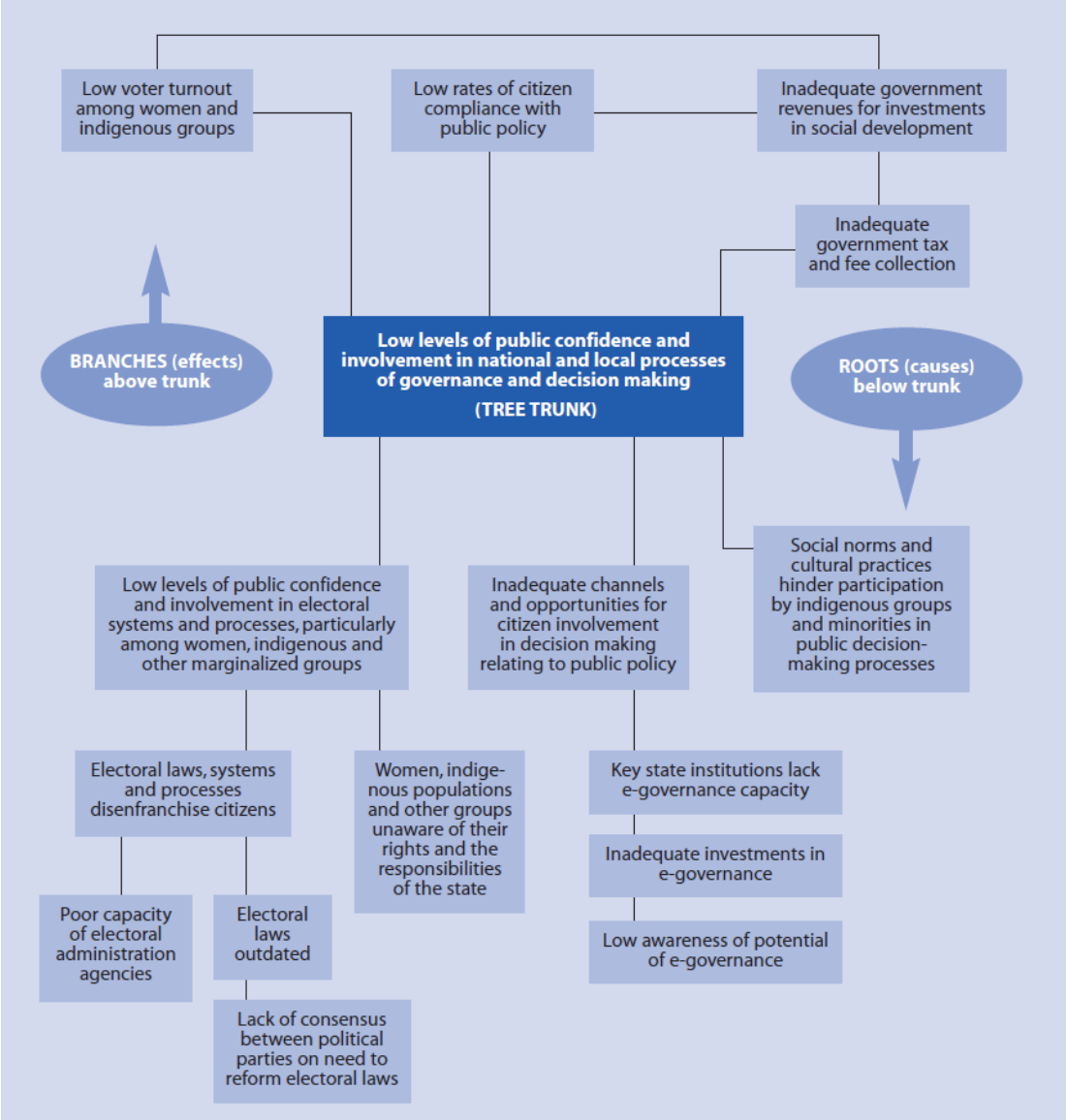
- Before discussing the design of an impact evaluation, it is important to understand how social or public programs are designed
- Designing a social program includes the following steps:
  1. Problem analysis
  2. Results map
  3. Analysis of assumptions, risks and unintended outcomes
  4. Results framework or theory of change
  5. Indicators

3.1 Problem analysis

- The goal is to study the root causes and major effects of problems in order to better design solutions
- Problem tree model:
  - Begin with a major issue or problem that was identified and write it down on the trunk of the problem true
  - Brainstorm on the major causes of the problem
  - Brainstorm the possible causes of the problem by asking "what is causing this to happen?"
  - Attach the answers to roots of the true
  - For each answer, drill down further by asking "why has this happened?"
  - Repeat this exercise for each cause identified
  - Once the roots of the tree are complete, check whether the tree provides a good understanding of what has caused the problem

- Effects of the problem:
  - Identify the most direct effects of the problem
  - Identify the main indirect effects of the problem
  - Discuss whether the problem affects groups in different ways

Example of a problem tree

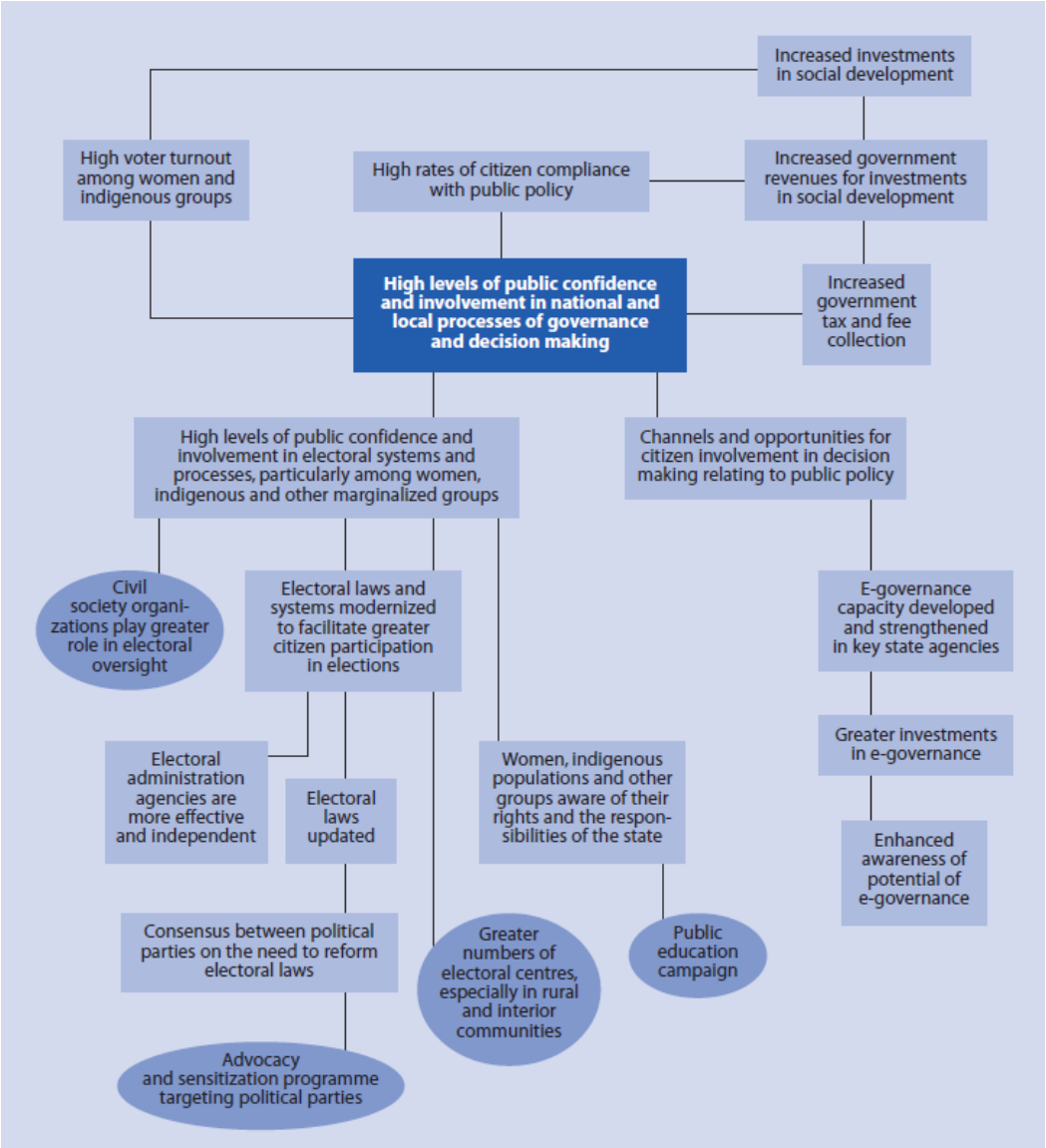


QUICK CHECKLIST FOR REVIEWING A PROBLEM TREE	YES	NO
✓ We have identified problems and causes that relate to the policy and legislative environment		
✓ We have identified problems and causes that relate to gaps in institutional capacities		
✓ We have identified problems and causes that relate to cultural and social norms		
✓ We have identified problems that affect men, women and marginalized populations, and the rights of different groups		
✓ We can see many layers of causes of the problems we have identified		
✓ We have defined the problems in the broadest terms, looking beyond the issues that individual agencies or stakeholders are concerned with		
✓ We have defined the problems and their causes without initially focusing only on the dimensions that one or more agencies have capacity to address through projects		

3.2 Results map/Theory of change

- A results map takes each problem identified on the trunk of the problem tree and reword it as the **immediate positive results** with long-term positive results or effects
- Map the complete set of lower level results:
  - Write down both the **immediate positive results** and all the longer term results of effects that they are trying to achieve
  - Work backwards and document the **major prerequisites** and changes needed for this result
  - Document other **lower level prerequisites** that are needed for the first set of changes and conditions to be in place
  - Results are not necessarily actions to be taken for stakeholders, but a set of **key things that must be in place**
  - Once intermediate changes have been identified, stakeholders should then **identify the interventions that are necessary to achieve them**

Results map


Increased investments in social development
High voter turnout among women and indigenous groups
High rates of citizen compliance with public policy
Increased government revenues for investments in social development
High levels of public confidence and involvement in national and local processes of governance and decision making
Increased government tax and fee collection
High levels of public confidence and involvement in electoral systems and processes, particularly among women, indigenous and other marginalized groups
Channels and opportunities for citizen involvement in decision making relating to public policy
Civil society organi-zations play greater role in electoral oversight
Electoral laws and systems modernized to facilitate greater citizen participation in elections
E-governance capacity developed and strengthened in key state agencies
Electoral administration agencies are more effective and independent
Electoral laws updated
Women, indigenous populations and other groups aware of their rights and the respon-sibilities of the state
Greater investments in e-governance
Enhanced awareness of potential of e-governance
Consensus between political parties on the need to reform electoral laws
Greater numbers of electoral centres, especially in rural and interior communities
Public education campaign
Advocacy and sensitization programme targeting political parties

QUICK CHECKLIST FOR REVIEWING A RESULTS MAP	YES	NO
✓ We have identified results that relate to addressing policy and legislative constraints		
✓ We have identified results that relate to addressing gaps in institutional capacities		
✓ We have identified results that relate to addressing relevant cultural and social norms		
✓ We have identified results to improve the condition of men, women and marginalized groups		
✓ We have identified results that address the rights of different groups in society		
✓ We can see many layers of results		
✓ We have defined the results in broad terms, looking beyond the specific contribution of individual agencies or stakeholders		
✓ The results map provides us with a picture of the broad range of actions that will be needed (including advocacy and soft support) and does not only focus on projects or tangible outputs		
✓ The results map shows us where action will be needed by both partners and non-partners in our effort		

3.3 Assumptions, risks and unintended outcomes

- In order to achieve the expected impact, there are complementary changes that are beyond the control of the program:
  - Assumptions are required in order the impacts to take place
  - Sometimes good intentions may lead to negative results
  - There may be risks that could prevent the planned results to be achieved

Assumptions

- Assumptions are the necessary and positive conditions that allow for a successful cause-and-effect relationship between different levels of results
- When thinking about positive changes, it is critical to ask the following: "if we say that having X in place should lead to Y, what are we assuming?"
- Assumptions may relate to the context within which stakeholders will works toward the desired result or refer to external factors about the behavior of other agents
- Assumptions related to lowest levels of map results are typically under control of the implementing agency

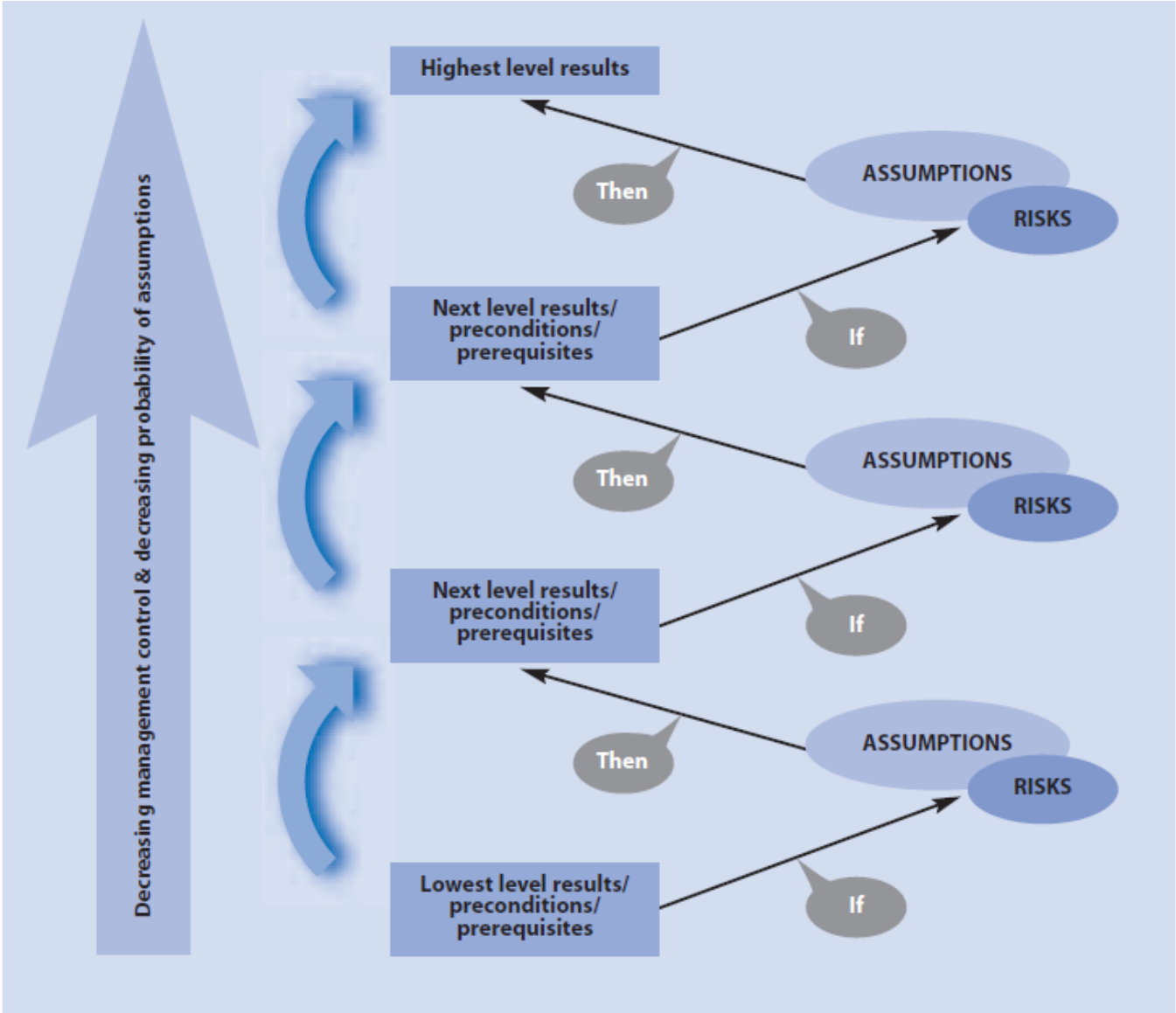
Risks

- Risks are potential events or occurrences beyond the control of the program that could adversely effect the achievement of results
- Steps are needed to mitigate their effects, depending on their likelihood, costs and impact
- How do they differ from assumptions?
  - Assumptions relates to conditions that should be in place for a program to go ahead, where these conditions have a high probability to happen
  - Risks relate to the possibility of external negative events that could jeopardize the success of the program

Unintended consequences

- These are another form of risk, but do not jeopardize the probability of the program's activities to happen. Rather, it refers to the possibility of the program taking place but leading to undesirable results
- Once assumptions and risks are taking into account, stakeholders should discuss and document any possible unintended results or consequences
- Stakeholders can device strategies to minimize the occurrence and impact of unintended outcomes

Assumptions and risks

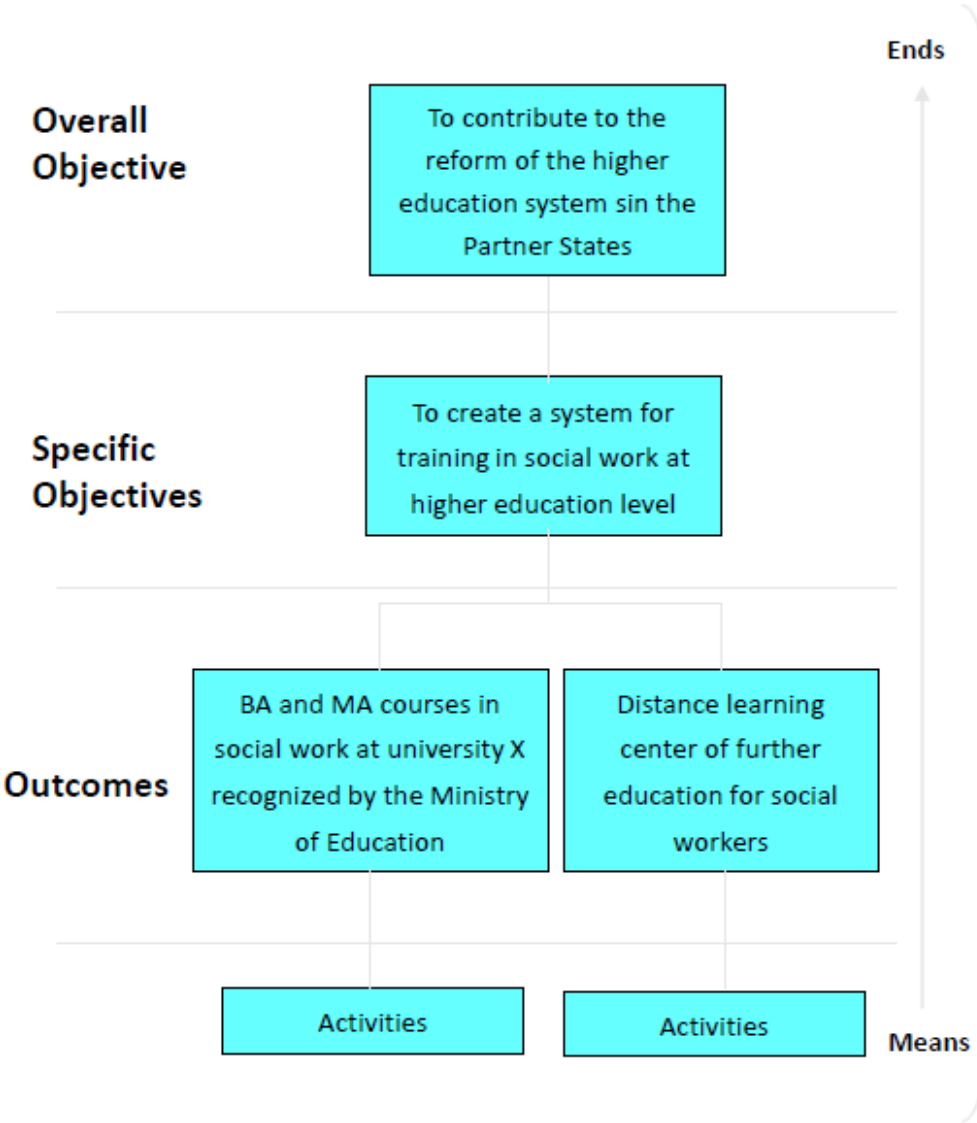


QUICK CHECKLIST FOR VALIDATING ASSUMPTIONS AND RISKS	YES	NO
✓ The assumed condition is outside the control of the programme or project		
✓ The assumed condition is necessary for programme success.		
✓ The assumed condition is **not** a result that could be included in the results framework		
✓ There is a **high probability** that the assumption will hold true		
✓ The assumption is specific and verifiable—its status can be checked by calling partners or donors		
✓ The assumption is stated as if it is actually the case		
✓ The risk is clearly beyond the control of the programme		
✓ The risk is NOT simply the negative restating of an assumption		
✓ The consequences of the risk are sufficiently grave as to pose a serious threat to overall programme success		
✓ There is a **moderate to high probability** that the risk may occur		

3.4 Results framework

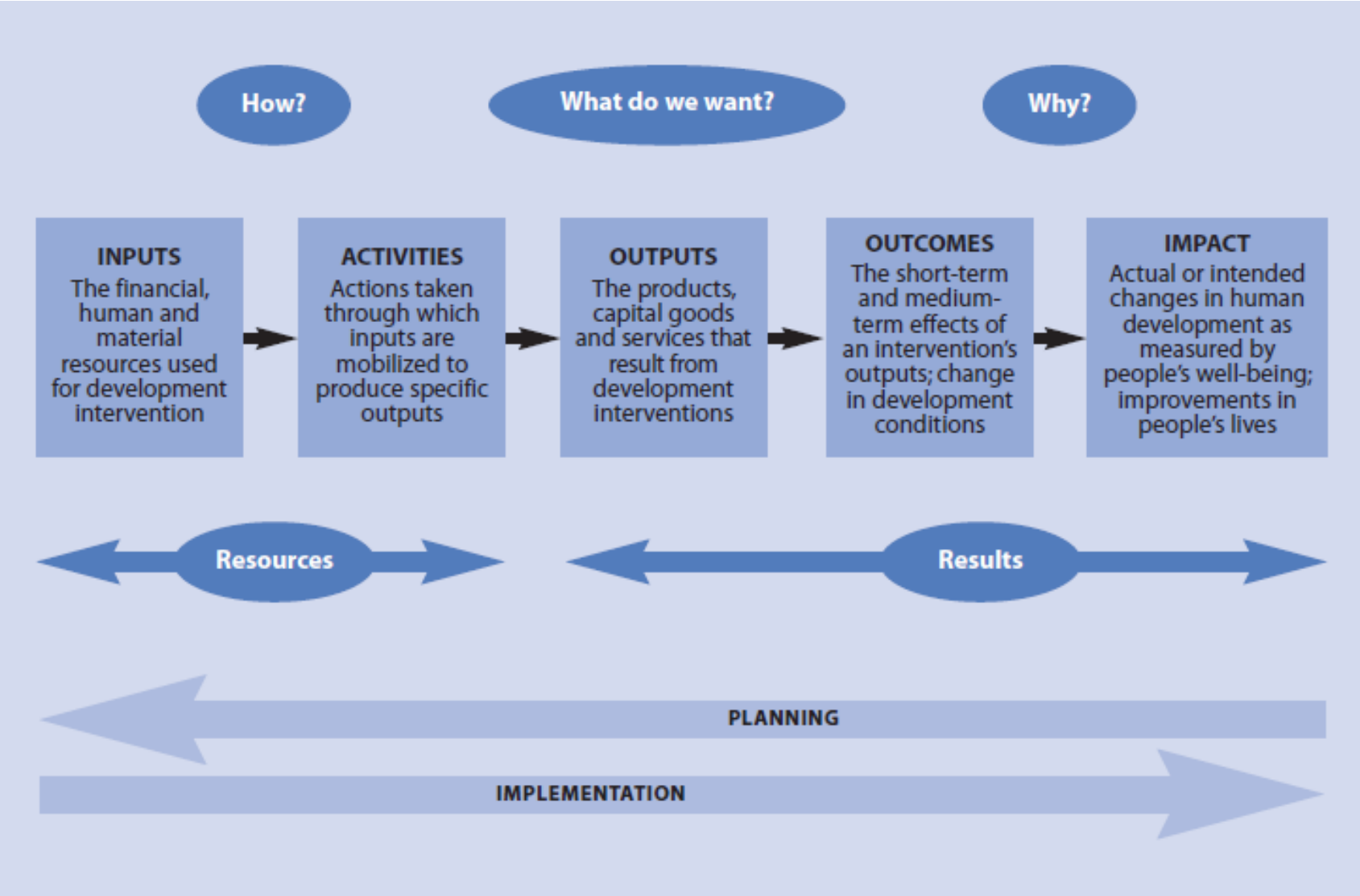
- The result map/theory of change is converted into the results framework (also known as logical framework)
- The logical framework incorporates the information from the results map and the analysis of assumptions, risks and unintended consequences and complements them with indicators and means of verification
- The results map is converted into a **results chain**

Converting the results map into the results chain



	Narrative
Overall Objective	To contribute to the reform of the higher education system sin the Partner States
Specific Objectives	To create a system for training in social work at higher education level
Outcomes	• BA and MA courses in social work at university X recognized by the Ministry of Education. • Distance learning center of further education for social workers.
Activities	Activities...

Results chain



Results framework

Results	Indicators	Baseline	Target	Means of Verification	Risks & Assumptions
Impact statement *(Ultimate benefits for target population)*	Measure of progress against impact				Assumptions made from outcome to impact. Risks that impact will not be achieved.
Outcome statement *(Short- to medium-term change in development situation)*	Measure of progress against outcome				Assumptions made from outputs to outcome. Risks that outcome will not be achieved.
Outputs *(Products and services—tangible and intangible—delivered or provided)*	Measure of progress against output				Assumptions made from activities to outputs. Risks that outputs may not be produced.
Activities *(Tasks undertaken in order to produce research outputs)*	Milestones or key targets for production of outputs				Preconditions for implementation of activities.

3.5 Indicators

- Indicators are signposts of change along the path to development. They describe the way to track intended results and are critical for monitoring and evaluation
- Indicators can help to:
  - Inform decision making for project management
  - Measure progress and achievements
  - Clarify consistency between activities, outputs, outcomes and impacts
  - Ensure legitimacy and accountability to all stakeholders
  - Assess project and staff performance

SMART indicators

S	**Specific:** Impacts and outcomes and outputs must use change language—they must describe a specific future condition
M	**Measurable:** Results, whether quantitative or qualitative, must have measurable indicators, making it possible to assess whether they were achieved or not
A	**Achievable:** Results must be within the capacity of the partners to achieve
R	**Relevant:** Results must make a contribution to selected priorities of the national development framework
T	**Time-bound:** Results are never open-ended—there is an expected date of accomplishment

Impact indicators

Sample Impacts	Sample Indicators (i.e., "What can we see to know if change is happening?")
■ Increased public participation in national and local elections, particularly by women, indigenous populations and other traditionally marginalized groups	■ Overall proportion of eligible voters who vote in the national (or local) elections ■ Percentage of eligible women who vote in the elections ■ Percentage of eligible indigenous people who vote in elections
■ Improved educational performance of students in region of the country	■ Percentage of students completing primary schooling ■ Pass rates in standardized student tests
■ Reduction in poverty and hunger	■ Poverty rate ■ Gini coefficient ■ Percentage of population living in extreme poverty ■ Level of infant malnutrition
■ People are healthier and live longer	■ Longevity ■ Infant mortality ■ HIV/AIDS prevalence rate

Indicators, baselines and targets

Indicator	Baseline	Target
IMPACT: Increased public participation in national and local elections, particularly by women, indigenous populations and other traditionally marginalized groups		
Overall proportion of eligible voters who vote in the national (or local) elections	2006: 42% of eligible voters voted in national elections	2010: 70% of eligible voters vote in national elections
Percentage of eligible women who vote in the elections	2006: 0% voted (women were not allowed to vote)	2010: 50% of eligible women vote in national elections
Percentage of eligible indigenous people who vote in elections	2006: 15% voted (no efforts were made to encourage or support voting by indigenous people living in the interior)	2010: 45% of eligible indigenous persons vote in national elections
OUTCOME: Electoral administrative policies and systems reformed to ensure freer and fairer elections and to facilitate participation by marginalized groups		
Percentage of public that believe that the electoral management process is free and fair	2006: 30% (based on last survey conducted)	2010: 80%
Percentage of women and minorities aware of their rights under the new electoral administration laws	2007: 20% of minorities said they were aware of their rights (survey done by [specify] agency; note: women were not allowed to vote)	2010: 70% of women and minorities aware of their rights
Percentage increase in number of women registered to vote	2007: 0% of women registered to vote (women were not allowed to vote)	20% annual increase in percentage of eligible women registered to vote
Percentage increase in number of indigenous people registered to vote	2007: 30% of eligible minorities registered to vote	20% annual increase in percentage of eligible minorities registered to vote
Ratio of voter registration centres per population in rural areas	2006: 1 centre to 11,000 people	2010: 1 centre to 4,000 people

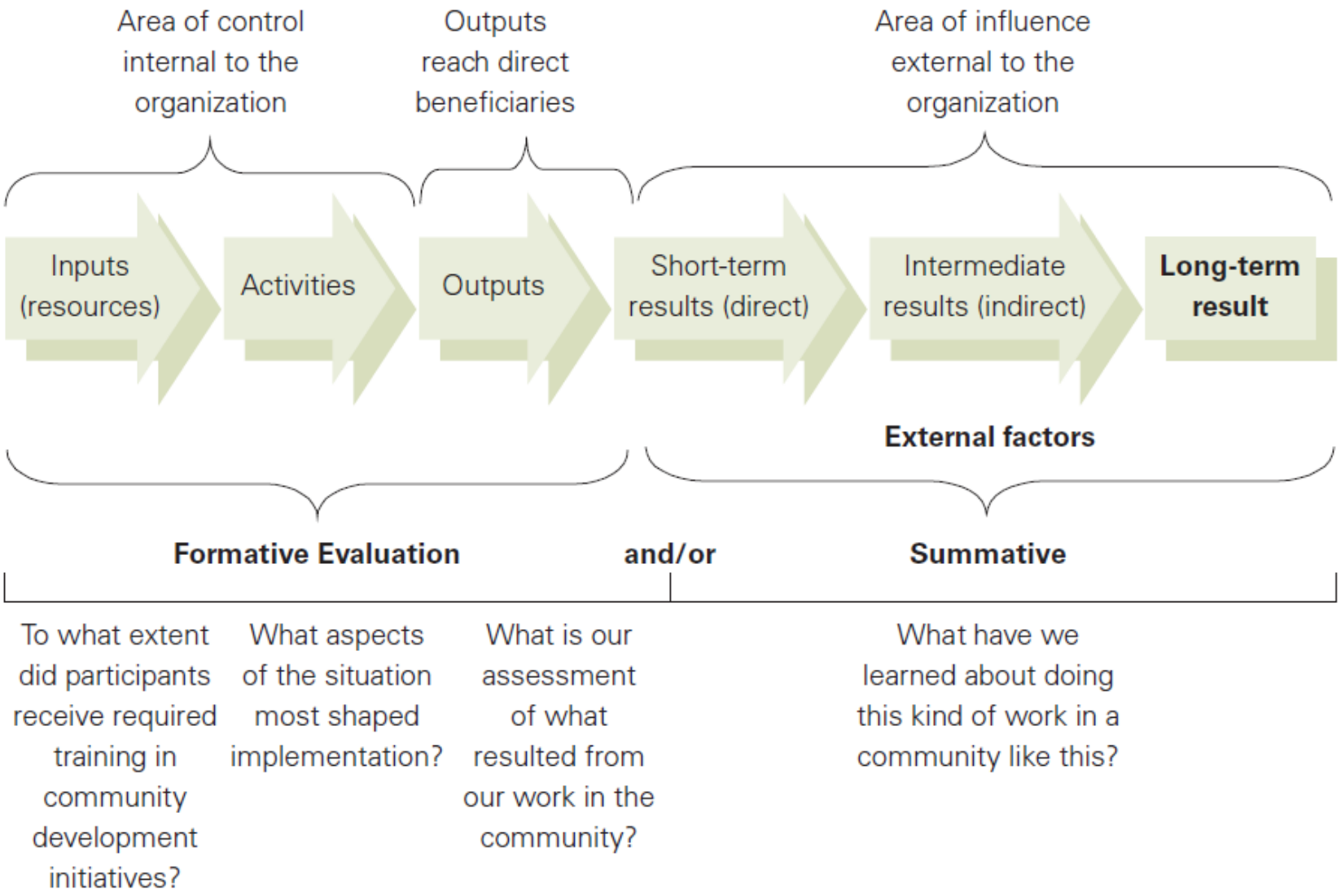
Sample results framework

Indicator	Baseline	Target	Means of Verification
IMPACT: Increased public participation in national and local elections, particularly by women, indigenous populations and other traditionally marginalized groups			
Overall proportion of eligible voters who vote in the national (or local) elections	2006: 42% of eligible voters voted in national elections	2010: 70% of eligible voters vote in national elections	Office of Electoral Administration's final report on elections
OUTCOME: Electoral administrative policies and systems reformed to ensure freer and fairer elections and to facilitate participation by marginalized groups			
Percentage of public that believe that the electoral management process is free and fair	2006: 30% (based on last survey conducted)	2010: 80%	Special survey to be undertaken as part of the electoral assistance project in 2008 and 2010
Percentage increase in number of women registered to vote	2007: 0% of women registered to vote (women were not allowed to vote)	2010: 20% annual increase in percentage of eligible women registered to vote	Office of Electoral Administration's database
Ratio of voter registration centres per population in rural areas	2006: 1 centre to 11,000 people	2010: 1 centre to 4,000 people	To be computed based on number of centres (Electoral Office database) in relation to population in rural areas (National Planning Agency's 2010 demographic survey)
OUTPUT 1: Draft new policy on electoral reform formulated and submitted to Cabinet			
Progress made in drafting new policy	2008: Agreement reached between major political parties on need to redraft electoral legislation	2009: 5 major public consultations held and white paper prepared on new policy	Report from government agency organizing workshops Record of Parliamentary proceedings (for submission of white paper) to be obtained from Office of Public Sector Information
OUTPUT 2: National electoral management agency has systems, procedures and competencies to administer free and fair elections			
Percentage of electoral centres using multiple forms of voter identification measures	2006: 0% of centres used multiple forms of voter identification	2009: 70% of centres use two or more forms of voter identification, including fingerprint identification (annual targets may be set)	Electoral Office database

4. Evaluation questions

- Why choosing evaluation questions is so important?
  - It gives direction to the evaluation and the selected evaluation design
  - Helps key stakeholders improve efforts, make decisions, and provide information to the public
- Program's theory of change provides a starting point to frame evaluation questions
- Type of evaluation questions:
  - Descriptive
  - Normative
  - Cause-effect

Theory of change and evaluation questions



Type of questions

Providing Free Measles Immunization

Intervention: Family clinics provide free immunization against measles to all children under the age of five in three regions of the country in one year.

Evaluation questions:

1. How did the clinics reach parents to inform them about the free immunization for their children? (descriptive question)
2. Did the program meet the target of providing immunization against measles to 100 percent of all children under the age of five in the three regions last year? (normative question)
3. Did the program use innovative methods to reach the children most at risk? (descriptive question)
4. Did the proportion of children contracting measles decrease as a result of the program? (cause-and-effect question)
5. Has there been a decline in child mortality from measles-related complications as a result of this program? (cause-and-effect question)

How to choose evaluation questions?

	Evaluation question									
Would the evaluation question	1	2	3	4	5	6	7	8	9	10
be of interest to key audiences?										
reduce present uncertainty?										
yield important information?										
be of continuing (not fleeting) interest?										
be critical to the study's scope and comprehensiveness?										
have an impact on the course of events?										
be answerable given the financial and human resources, time, methods, and technology available?										

Evaluation design matrix

Design matrix for: _____

Main evaluation issue: _____ **General evaluation approach:** _____

Question	*Subquestion*	*Type of subquestion*	*Measure or indicator*	*Target or standard (normative)*	*Baseline data?*	*Data source*	*Design*	*Sample or census*	*Data collection instrument*	*Data analysis*	*Comments*

Example Design matrix

Main Evaluation Issue: Should this program be reauthorized? **General Approach:** Quasi-Experimental Impact Evaluation

Questions	Subquestions	Type of (sub) question	Measures or indicators	Target or standard (if normative)	Baseline data?
1. What services did the program provide to whom?	**1.A.1.** In what vocational skill areas were participants trained?	Descriptive	Vocational skill areas program offered to trainees for certification	NA	NA
	1.A.2. Were there changes over time in participation by skill area?	Descriptive	Same as above by number of participants each year NA	NA	Yes
	1.B.1. What support services were offered by the program?	Descriptive	Support services (e.g. literacy, counseling) offered by the program	NA	NA
	1.B.2. What proportion of participants received support services?	Descriptive	Number and percent of trainees receiving each type of support	NA	NA
	1.C. What were the most popular certification areas selected by trainees in the vocational training program?	Descriptive	Number of trainees by certificate areas	NA	NA
	1.D. To what extent did the program provide certification in areas forecast as high demand for the next 5-10 years?	Descriptive	List of vocational areas forecast as high demand over the next 5-10 years	NA	NA
2. To what extent was there gender equity in the services delivered?	**2.A.1.** Did equal numbers of males and females participate in the program?	Normative	Number and proportion males/females receiving vocational training	Program authorizing documents indicate 50% participation goal for females	NA
	2.A.2 Is receipt of support services related to gender?	Normative	Proportions by gender receiving each of the support services offered	Program authorizing documents indicate gender equality a goal	NA

Example Design matrix...

Design	Data sources	Sample	Data collection instrument	Data analysis	Comments
1.A.1. One Shot	Program records (MIS) Program Director	For each of past 5 years	Record Retrieval Document 1 Program Officials Interview Guide	Frequency count Content analysis	Data sources should match; note any discrepancy and explain
1.A.2. Time Series	Same as above	Same as above	Same as above	Same as above by year	Graphic would be good here
1.B.1. One Shot	Program records (MIS)	Census over past 5 years	Record Retrieval Document 2	List	Check for duplicates such as M. Smith and Mary Smith
1.B.2. Same as above	Same as above	Same as above	Same as above	Frequency count	Note that participants can receive more than one support service
1.C. One Shot	Program records (MIS)	Census over past 5 years	Record Retrieval Document 2	Frequency count	Graphic
1.D. Time Series	Labor Ministry Annual Reports on Short-, Mid-, and Long-Term Labor projections	Reports for each of past 5 years	Record Retrieval Document 3	Trend analysis and forecast for each certification area offered over the past 5 years	Note changes in trends and program's responsiveness to them Note whether there were potential growth areas in which the program did not offer training
2.A.1. Time Series	Program records (MIS)	For each of past 5 years	Record Retrieval Document 1	Frequency counts by gender; present as line chart so trend over 5 years is clear. Compare to standard	Show standard in heavy black line across the line chart so it is easy to see.
2.A.2. Same as above	Same as above	Same as above	Same as above	Same as above	Note if there were changes over time

Example Design matrix...

Questions	Subquestions	Type of subquestion	Measures or indicators	Target or standard (if normative)	Baseline data?
3. Was the program effective?	**3.A.** To what extent were the annual job placement targets met or exceeded?	Normative	Job placement rates by year	Yes. 80% of those completing the program	NA
	3.B. To what extent were the annual average placement wage rates met or exceeded?	Normative	Job placement wages for each year	Yes. $2 per hour years 1-3 and $3 an hour years 4&5	NA
	3.C. To what extent were participants placed in jobs that matched their certification areas?	Descriptive	Trainee certificate area and job placement area	Implicit standard only so treated as descriptive question	NA
	3.D, What was the program's drop-out rate?	Normative	Number entering the program each year and number graduating each year	Program documents indicate should not be more than 10%	NA
4. Was the program cost-efficient?	**4.A.** Was the program cost per participant reasonable in relation to similar programs?	Descriptive	Cost per placed trainee compared to other similar training programs	Implicit standard only so treated as descriptive question	NA
5. To what extent was instructor turnover a problem?	**5.A,** What was the turnover rate of instructors?	Descriptive	Turnover rate annually and for instructors	None set. Implicit standard is that it should be low	NA
	5.B. How long did instructor positions stay vacant?	Descriptive	Average length of vacancies and range	None set. Implicit standard is that it should be low	NA
	5.C. Were equally qualified instructors found as replacements?	Descriptive	Years teaching experience Years working in area of certification	None set. Implicit standard is that they should be comparable	NA

Example Design matrix...

Design	Data sources	Sample	Data collection instrument	Data analysis	Comments
3.A. Time Series	Trainee records by year for each of 5 years (MIS)	Census of those placed by year	Record Retrieval Form 4	Comparison to standard each year and cumulatively across the 5 years.	Need to validate the information in the records by confirming placements and starting rates with employers for a sample of trainees. Recall will likely be a problem for past two years.
One Shot	Trainee records for the last year (MIS) Employers/ Employers' records	Random sample	Employer Interview Guide Employer Record Form 1	Match of MIS info with employer information	
3.B. Time Series	Trainee records by year for each of 5 years	Census of those placed by year	Record Retrieval Form 4	Comparison to standard each year and cumulatively across the 5 years.	Need to validate the information in the records by confirming placements and starting rates with employers for a sample of trainees. Recall will likely be a problem for past two years.
One Shot	Trainee records for the last two years (MIS) Employers/ Employers' records	Random sample	Employer Interview Guide Employer Record Form i	Match of MIS info with employer information	
3.C. One Shot	Trainee records for past 2 years showing certification area (MIS)	Census of those placed	Record Retrieval Form 4	Frequency count of matches by year and cumulatively	
	Employers hiring trainees over past 2 years	Random sample	Employer Interview Guide Employer Record Form 1		
3.D. Time Series	Trainee records by year for each of 5 years (MIS)	Census	Record Retrieval Form 4	Comparison to standard each year and cumulatively	
4. One Shot	Program Financial Officer Placement rates (see 3A) Program Financial Statements	Census- all 5 years of financial statements	Interviews	Cost per participant Cost per participant placed	Hope to be able to compare cost per placed trainee to that of other similar training programs.
	Existing evaluations of similar training programs		Literature review	Content analysis	
5.A. Time Series	Program employment records	All 5 years	Record Retrieval Form 5	Frequency counts by year, range and average	
5.B. Time Series	Program financial records	All 5 years	Record Retrieval Form 5	Frequency counts, range and average	
5.C. Time Series	Program employment records- c.v.s	All 5 years	Record Retrieval Form 6	Comparisons of staff c.v.s Across the 5 years	

Example Design matrix...

Questions	Subquestions	Type of (sub) question	Measures or indicators	Target or standard (if normative)	Baseline data?
6. To what extent were trainee dropouts a concern?	**6.A**. What were the numbers and percentages of males and females dropping out of the program each year?	Descriptive Proportions by gender	Number and proportion males/ females starting the program by year and dropout rates by year.	Background documents indicate 10% is acceptable; Implicit that it would be the same rate for each gender	NA
	6.B. What were the certification areas from which they dropped?	Descriptive	Above by certification area.	NA	NA
	6.C. What were the common reasons for dropping out of the program?	Descriptive	Most frequent reasons for dropping out.	NA	NA
	6.D. How concerned were program officials about drop-out rates?	Descriptive	Awareness of drop-out rates Opinion on whether problem Actions taken	None specified.	NA
7. To what extent do those trainees placed in jobs earn more than they would have absent the training program?	**7.A**. What are the job retention, salary increase, promotion, and firing rates of placed participants compared with others similar in characteristics who dropped out of the training program?	Cause & Effect	Placed participant job retention, starting salary, salary increases, and promotion and firing rates over two years compared with (i) others hired by the firms for similar positions over a comparable period; (ii) pre-program earnings; (iii) earnings of program drop-outs.	None specified	Yes, on placement wages
	7.B. What are employers' views of the performance of placed employees compared with others hired with similar characteristics who did not receive the training program?	Cause & Effect	(i)Likelihood of hiring training participants absent the program and (ii) hiring more program grads; (iii) views on job performance compared with others.	None specified	Initial placement wages; pre-training wage, if any

Example Design matrix...

Design	Data sources	Sample	Data collection instrument	Data analysis	Comments
6.A. Time series	Trainee records by year for each of 5 years	Census by year	Record Retrieval Form 4	Frequency distribution of drop-outs by year; percent drop-out by year by gender, and totals	
6.B. Time series	Trainee records by year for each of 5 years	Census by year	Record Retrieval Form 4	Cross tab of certification areas by frequency of program drop-out annually and cumulatively	Might run test of significance—Chi Square?
6.C. One Shot	Trainee records for past two years	Census of program drop-outs for the two year period	Former participant survey	Frequency distribution of reasons for dropping out of program from (i) participant perspectives, and (ii) training program officials' views	Triangulate
	Training program officials	Senior officials	Program Officials Interview Guide		
6.D. One Shot	Trainee records for past two years Training program officials	Census of program drop-outs for the two year	Program Officials Interview Guide	Content Analysis and frequency counts	
7.A. Quasi-experimental Design- Non-equivalent Groups	Employers and Employer records	Census for prior two years	Employer Interview Guide Employer Record Form 1	Content analysis/ frequencies	No comparison group formed at program initiation; proxies need to be used
	Program drop-outs		Former Participant Survey		
7.B. One Shot	Employers and Employer records	Census for prior two years	Employer Interview Guide Employer Record Form 1	Content analyses/ frequencies	Note C&E design limitations; are confidential, no access