PAI 705
McPeak
Lecture 12

Evaluation Research

- Research undertaken for the purpose of determining the impact of a given intervention.
    - Has the intervention produced the intended results?
        - Use of social research procedures to systematically investigate the effectiveness of social intervention programs.
    - Used to study, appraise, and help improve programs.
    - So not just to say 'has it worked' but also at times 'what could be done to make it work better in the future'.
    - Also can be important to identify not just 'did it work' but also identify unintended consequences.
    - It is a rapidly growing field.

I am going to provide an overview first based on the chapter in the Babbie text, then we will revisit this topic with the World Bank discussion

Morra-Imas, L. G., Morra, L. G., & Rist, R. C. (2009). The road to results: Designing and conducting effective development evaluations. World Bank Publications. You can download it online here:
https://openknowledge.worldbank.org/bitstream/handle/10986/2699/526780PUB0Road101Official0Use0Only1.pdf?sequence=1

US federal demand for this research. US federal funding for this research.

- Evidence based policy making.
- https://results.usaid.gov/
- http://betterevaluation.org/resources/guide/dfid_appraoch_to_value_for_money

Topics for Evaluation Research:

- Needs assessments. Studies to determine the existence and extent of problems. Can be targeted to distinct subpopulations.
- Cost benefit studies. Looking at the present value of benefits compared to the present value of costs. Looking at flows over time and discounting to state in Net Present Value. Cost benefit ratio, IRR. Discounting is a key element.
    - MCC https://www.mcc.gov/our-impact/err

- Monitoring a steady flow of information about the object of the research, to report on the attainment and status of the activity. M&E. Are you doing the meetings you promised you were going to do.
- http://usaidprojectstarter.org/content/performance-monitoring-indicators

Table 1: Programme M&E Framework

Impact	Indicators	Base Situation	Expected End Point	Information needed	Methods/Data Collection	By whom	When
% of people with improved wellbeing in the seven cercles and départements in Mali and Senegal as a result of BRACED support. (disaggregated Cercles and Departments, by gender and age)	To be defined when relationship between aspects of resilience and wellbeing has been explored in this context.	PDP phase looked at aspects of wellbeing.	Expected to understand contribution to impact but not to see change in this timeframe.	Identify investments and determine likely linkages to well-being.	Resilience assessments, HH survey and family portraits help to establish links between resilience targeted by the investments and wellbeing. When this is identified secondary data will be tracked.	Country M&E officers.	Initial assessment at baseline (end of year 1). Then annual.
Outcome	**Indicators**	**Base Situation**	**Expected End Point**	**Information needed**	**Methods/Data Collection**	**By whom**	**When**
Vulnerable communities in 3 Circles and 4 Departments in Mali and Senegal improve their resilience to climate extremes	1. Improved resilience linked to support from DCF (disaggregated quantitative indicator KPI4)	Nil re: DCF support but other ongoing interventions may affect benchmark Baseline resilience assessment	Survey of individuals with improved resilience in relevant domains (*see KPI 4 methodological note – Annex 4*)	Understand community perception of climate change, stresses and resilience strategies employed Deeper analysis of climate challenges and intra family responses by agro zone & livelihood types	Resilience assessments Focus group discussions Household Surveys 600 households	DCF field staff w. Animateurs Survey enumerators	Baseline resilience assessment June – end Nov 2015; 7 localities Follow on 600 households *by Dec 2015.* Then annual (end of milestone 2 and final evaluation).
	2. Use of climate information, to inform adaptation responses	Use of climate information at local authority level low	All LA are making use of climate information to inform local	Type of information judged likely to be useful; capacity for	TAMD Scorecards Indicators 1 & 5 (use of climate information and	DCF Field Staff with support for analysis M&E staff	Annual.

			development planning	analysis and use of this within LA	integration into planning)		
Outputs	**Indicators**	**Base Situation**	**Expected End Point**	**Information needed**	**Methods/Data Collection**	**By whom**	**When**
1. Devolved finance and planning mechanisms are established and functional in three Cercles (Mali) and four départements (Senegal) to support community-prioritized investments in public goods that build climate resilience.	1.1 Number of devolved climate adaptation financial mechanisms established with local authority partners, transparently managed and fully accountable	Nil devolved climate funds. Unclear to date what part of other existing funds currently reach local level structures in target areas: 6 finance sources 6 / programme funds named in Mali; NIE expedites adaptation finance in Senegal[i]	Fund placed at region level Mopti with 3 Cercle specific funds operational (Mali) Fund potentially housed within PNDL at national level with 4 funds at department level (Senegal)	Lessons to be drawn from current dispersal of climate funds Agreements and delivery of these As DCF are operationalized, monitor agreed performance criteria such as: * Nos. of grants * Value of awards * Duration of cycle application to completed implementation	Ongoing survey & mapping of contexts MOU Internal monitoring system of public investments funded through DCF	DCF M&E staff Regional fund managers; Monitored by government audit process and DCF M&E Staff	Quarterly from start up Yr 1
	1.2 Number and type of public good investments responsive to community prioritisation, demonstrating social and gender inclusion	Range of strategies in use to plan locally for climate adaptations[ii] Assessment of women / youth concerns & inclusion in these mechanisms is made as part of in resilience assessment	Communities manage fund investments to implement projects. Community members, including those from more vulnerable categories, are supported within the range of public	Monitor number of projects proposed, implemented. Scope community priorities Monitor the range of public goods investments Assess how effectively full	Monitor finalization of projects. Monitor grants made against community mapping of priorities.	DCF M&E staff DCF staff lead commune level monitoring group to track locally	Quarterly from grant making start up: Yr2

			investment activities chosen	inclusion is delivered within responses selected by community, process to be established			
	1.3 Number of local governments using TAMD Track 1 as part of planning & Track 2 indicators as part of performance monitoring processes	Participatory capacity assessment of LA mandate and systems made using TAMD adapted scorecard – low capacity in several areas	LA improve capacity to monitor impact from funding public goods	What tools or processes used by LA to review development plans in general; climate adaptation specifically	TAMD Scorecards	DCF M&E and Field Staff	Quarterly from start (Q2)
	1.4 Proportion of LA with improved capacity for CRM	Output from participatory capacity assessment of LA provides basis for ranking	Local Authorities (LA) rate self improvement in their capacity to respond to CC	How do leaders and officials deliver against their responsibility and need to respond to CC	Capacity checklist TAMD indicators 2 to 4	DCF M&E and Field Staff	Quarterly from start (Q2)
	1.5 Number of individuals trained to manage the devolved climate adaptation financial mechanisms, that rate the trainings positively	Nil before trainings	A majority of participants rate the trainings positively	Assess the rate of participant's satisfaction concerning the trainings	Training assessment by each participant at the end of the sessions.	DCF M&E and Field Staff	Quarterly from start (Q2)
Outputs	**Indicators**	**Base Situation**	**Expected End Point**	**Information needed**	**Methods/Data Collection**	**By whom**	**When**
2. Vulnerable communities in three Cercles (Mali) and four Départements (Senegal) benefit	2.1 Number of people directly accessing initiatives funded by climate resilience	Nil by definition KPI 1 PDP report on knowledge and Past adaptation	Numbers of communities and groups from 8 communes (Mali) and 24 (Senegal) whose members	KP1 monitoring How many individuals access public investments funded via DCF	Baseline resilience assessments set range of community priorities	LA M&E system DCF staff	Quarterly from YR1

from public good investments that build resilience and reflect community priorities	investments in the Cercles and Départements (disaggregated by gender and age)	projects, notes poor considerations of gender with . involving women in community decision-making, problematic.	are direct beneficiaries of DCF public investments		Grant awards from regional fund Group level info on membership and impact	DCF staff & IIED facilitation	

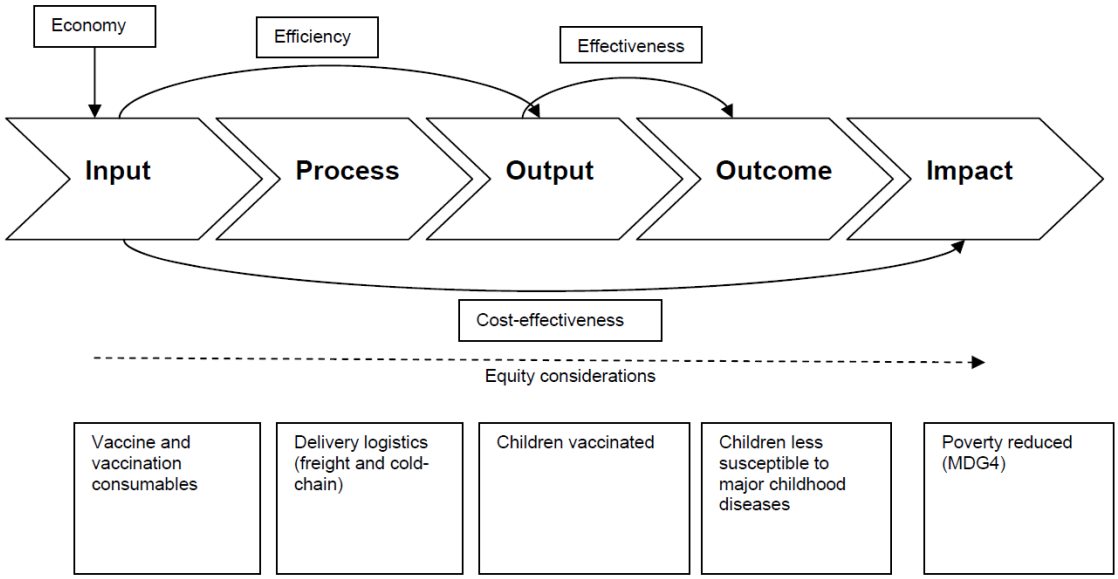
- Program evaluation, outcome assessment, impact evaluation.
    - Were people made better off by attending the meetings.
    - In what domains?
    - In the ways you predicted ex ante?
    - In ways that you did not predict ex ante?

https://usaidlandtenure.net/wp-content/uploads/2016/09/LAND-Oromia_Baseline-Report_2016-02.pdf

https://www.land-links.org/wp-content/uploads/2016/09/USAID_Land_Tenure_ERC_LAND_Afar_IE_Design_Report.pdf

Formulating the Problem.

- Issues of measurement.
    - What was the intended result of the intervention?
    - To what extent is it measurable?
    - To assess, we need to operationalize, observe, and recognize whether the intended result is present or absent following the intervention.
        - How can it be quantified?
        - Can it be converted into a cash value?
    - Can go at times to things like the mission statement. The donor intent. The project proposal.
    - This is where the program management class links up with the ideas of this course.



Specifying outcomes.

- What is the response variable? What were we trying to change or have an impact on?
- What are different ways of measuring this outcome? Consider different approaches and make sure the consumers of your evaluation product are in agreement both with your definition of the response variable and the means to operationalize the question.

Measuring experimental contexts.

- Not just the experiment has happened in the time under study, but other 'all else equal' factors have changed as well. Need to control for this. Classic response is to have

treatment and control. In addition, not just looking at the outcome in question but other conditioning variables that may explain differential impact within the treatment group if it exists.

- How do we measure the intervention? What was the stimulus, and how can we measure the degree of the stimulus? She went to one training, he went to three, control went to none….. The extent and quality of their participation in the trainings can be measured. What if the trainers differed, and some are more effective at conveying the skill set than others? You would want to track this.

Specifying the population.

- Who is the relevant population for the treatment? How closely matched to them are the population of the control? If the control is of the same population type, why were they not treated as well?

New versus existing measures.

- To what extent to you take questions used in other surveys for similar objectives? To what extent do you adapt them to the specific context and purposes of your objective?

PERCEPTIONS OF TENURE SECURITY

HOUSEHOLD/WIVES SURVEYS

FARMLAND

In the next 1-2 years, how likely is it that someone from within your extended family will take over the use of this field/plot without your household's permission/agreement?	1=Very Likely 2=Likely 3=Neutral 4=Somewhat unlikely 5=Very unlikely 888=Don't know 999=Prefer not to reply	PRADD, TGCC, CFP
In the next 5 years, how likely is it that someone from within your extended family will take over the use of this field/plot without your household's permission/agreement?	1=Very Likely 2=Likely 3=Neutral 4=Somewhat unlikely 5=Very unlikely 888=Don't know 999=Prefer not to reply	PRADD, TGCC
In the next 1-2 years, how likely do you think it is that the local government authorities will take over the use of this field/plot without your household's permission?	1=Very Likely 2=Likely 3=Neutral 4=Somewhat unlikely 5=Very unlikely 888=Don't know 999=Prefer not to reply	PRADD
In the next 5 years, how likely do you think it is that the local government authorities will take over the use of this field/plot without your household's permission?	1=Very Likely 2=Likely 3=Neutral 4=Somewhat unlikely 5=Very unlikely 888=Don't know 999=Prefer not to reply	PRADD
In the next 1-2 years, how likely do you think it is that private investors will take over the use of this field/plot without your household's permission?	1=Very Likely 2=Likely 3=Neutral 4=Somewhat unlikely 5=Very unlikely 888=Don't know 999=Prefer not to reply	PRADD, CFP

In the next 5 years, how likely do you think it is that private investors will take over the use of this field/plot without your household's permission?	1=Very Likely 2=Likely 3=Neutral 4=Somewhat unlikely 5=Very unlikely 888=Don't know 999=Prefer not to reply	PRADD
The boundaries of my {farmland} are clear and respected by people in this village.	1=Strongly Agree 2=Agree 3=Neutral 4=Disagree 5=Strongly Disagree 888=Don't Know 999=Refuse to answer	PRADD
I am confident that a government/investor cannot take any of my {farmland} without negotiation and fair compensation.	1=Strongly Agree 2=Agree 3=Neutral 4=Disagree 5=Strongly Disagree 888=Don't Know 999=Refuse to answer	PRADD
In the next 3 years, how likely do you think it is that people from a neighboring village will encroach/cross-over to use this field?	1=Impossible/would never happen 2=Highly Unlikely 3=Unlikely 4=Unsure/Don't know 5=Likely 6=Highly Likely 7=Happening right now 888=Don't know 999=Prefer not to respond	CFP, HH and Wives
How likely do you think it is that other households within your village may try to cross-over your boundaries (step on your side) and take or use some of this field in the next 1-3 years?	1=Very Likely 2=Likely 3=Neutral 4=Somewhat unlikely 5=Very unlikely 888=Don't know 999=Prefer not to reply	TGCC, CFP
How likely do you think it is that other households within your village may try to crossover your boundaries (step on your side) and take or use some of this field beyond 4 years from now?	1=Very Likely 2=Likely 3=Neutral 4=Somewhat unlikely 5=Very unlikely 888=Don't know 999=Prefer not to reply	TGCC
How likely do you think it is that elites/big people may take this field without your household's permission/agreement in the next 1-3 years?	1=Very Likely 2=Likely 3=Neutral 4=Somewhat unlikely 5=Very unlikely 888=Don't know 999=Prefer not to reply	TGCC, CFP, CLPP

Operationalizing Success / Failure.

- One basic standard is classic cost benefit analysis for an efficiency answer to the question. Do the discounted sum of benefits outweigh the discounted sum of costs?
- But what if the outcome is not easily translated into cash terms? The training improved children's test scores by 15%. So what is that worth? At times, you could look at different programs and consider the least cost way of making a given increase in test scores happen?
- As something that is obvious, but potentially overlooked, make sure you include indicators of what the planning documents of the program said they were trying to do. You need to orient yourself to the planning documents of the program to make sure you understand and measure what they said they were trying to accomplish.

Types of Evaluation Research Designs.

Experimental Designs.

Assign subjects randomly to a treatment group and a control group.

Survey baseline.

Apply stimulus.

Survey repeat.

A classic example:

http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0262.2004.00481.x/abstract

Quasi-experimental designs. Not a pure experiment, but there is something in the nature of it that you can make it 'like' an experiment.

Time series. Return to the idea of 'before' and 'after' and the contrast to 'with' and 'without'.

Add in the idea of a 'nonequivalent' control group. This is a control that is not part of the study. You try to make it 'similar' to the treated population. It is not created as a random sample. It is deliberately selected to be like your treatment sample.

Multiple time series designs. Use more than one set of time series data and contrast the patterns seen in the different data sets. There should be something comparable that makes contrasting the patterns informative.

Figure 12-3.

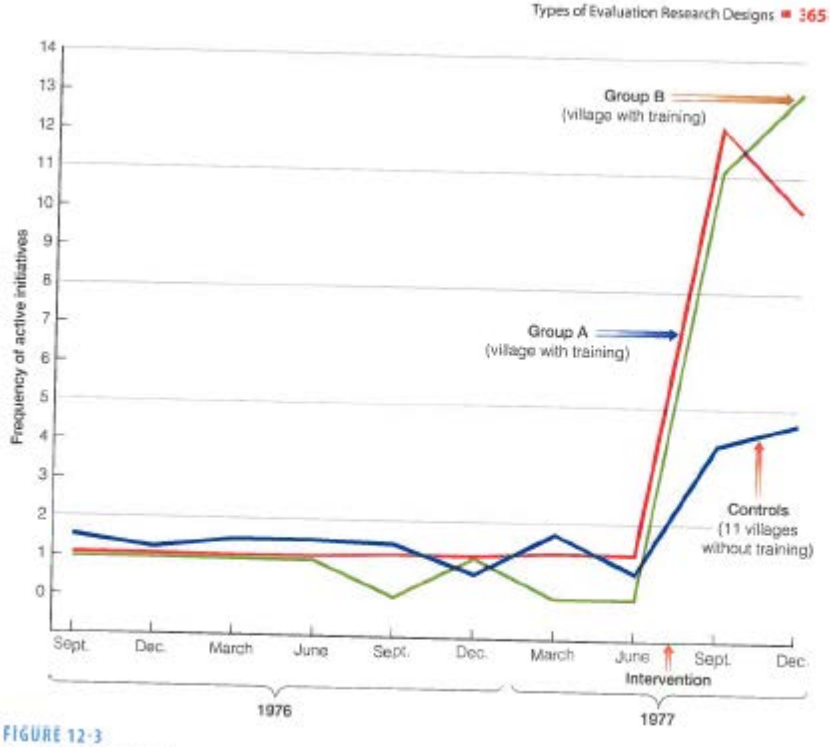


FIGURE 12-3
Active Initiatives over Time
Source: Rajesh Tandon and L. Dave Brown, 1981, "Organization-Building for Rural Development: An Experiment in India," *Journal of Applied Behavioral Science* (April–June): 182.

Note how hard it is to avoid 'contagion'.

Qualitative evaluation. Sometimes talking to people in depth gives you insight into things you would not get otherwise. The narrative explains things and how they connect.

Quantitative and qualitative often connect and complement each other. You find a pattern in the data from the quantitative side and you interpret it through a qualitative discussion.

Problems encountered in evaluation research.

Logistical problems.

- Nobody has a population list from which to sample.
- People who were on the list moved away or died.
- People refuse to cooperate.
- You can't get permission to run the survey from local authorities.
- Your starting sample size was fine but by the time you get to the evaluation you have too few.
- People don't want to be control, they want to be treatment.
- Up front, you are not sure where the project will be implemented.
    - Up front, you are not sure what is the nature of the project to be implemented.

At a basic level, people don't like getting evaluated. It is not necessarily in their interest.

- Authorities may send all the lowest performers your way to undermine your research / for their own reasons (relatives).
- Authorities may send all the highest performers your way to influence your findings / for their own reasons (graduation).

Uses of evaluation findings.

Clearly, to identify success and learn.

But also keep in mind, people have agendas.

The findings will influence what is expanded, reduced, terminated.

Officials, NGO workers, all kinds of people have agendas that are going to be influenced by the findings / have reasons to embrace, selectively interpret, ignore, undermine, sabotage.

One other area to consider is Social Indicators Research.

Monitoring aggregate statistics that are reported at the population level. Birth rates, death rates, total fertility rates, maternal mortality rates…..

Increasing role of computer simulation. Can start building models with larger data sets and simulate interventions and outcomes. Appeal to large quantitative data and do *ex ante* simulation to predict impacts.

http://www.awhere.com/

https://esoko.com/

Ex ante impact evaluation as a larger idea; survey people who know something about a topic to see what they think will have the biggest impact. Can use this to prioritize and allocate funding.

Ethical issues of Impact Evaluation

The treatment may be controversial.

Omission of treatment to control population is problematic.

There may be pressure on 'treatment' sample members.

They may not have an opt out option if the treatment is at a community level (think of worms and school selection).

Biased impact evaluations may fit into somebody's agenda