Meeting 3

Process Evaluation
Process Evaluation Examples
Formative Evaluation
Program Monitoring
Implementation Analysis

20170918

Example

- An idealistic young (ok, not so young) researcher wants to examine the impact of financial incentives on educational performance.

- Based on the literature he observes that financial incentives don't always seem to work.

- A colleague, a psychologist, observes that money changes motivation from intrinsic to external. And that unless students have right kind of self-belief, incentives can demotivate them.

A school intervention

- The researchers create a treatment to teach students to "believe in themselves" and then offer them prizes to see if they worker harder and achieve more.

- An evaluation later shows the program did succeed in raising attendance (the proximate goal) but not actual achievement.

- Why did the treatment fail?

Process evaluation

- Describes how a program is operating
- Assesses how well it performs intended functions
  - Built on program process theory
- "Verifies what the program is and whether or not it is delivered as intended to the targeted recipients" (Scheirer 1994)
  - As intended = program theory is one way of determining

Why is process evaluation important?

- Tells us a lot about the program
  - Replication – if don't know what it is, can't repeat
  - Generalizability – what is it that produces outcomes
- Pre-requisite for meaningful outcome evaluation
  - Explain results
  - Preventive: If no program or bad program, not worth evaluating
- Illuminates causal model
  - Link between program theory and outcome evaluation
    - Describes program activities and outputs
    - Establishes that these happened
    - That the underlying assumptions hold up
  - Helps identify what went wrong
    - Flawed implementation
    - Flawed Program Process Theory
      - Or even Flawed Program Impact Theory
- And recall: "program" can be any variation here, e.g., a shift in a laws.

Purposes of process evaluation

- Planning outcome evaluation
  - Does intervention exist at all?
  - Is intervention ready for impact evaluation? (Evaluability)
    - Fully or partially implemented (strong or weak test)?
    - Stable
    - Has fidelity to the model
- Helping program
  - Formative evaluation (program monitoring)
    - Feedback to fine tune
  - Official documentation of the program
- Understanding research results
  - How and/or why program worked or didn't work
  - Tease out which aspects worked and which didn't
  - Refine or test components of the logic / causal model
  - Determine how replicable
    - Feasibility of implementation
    - Contextual features/factors that may require adapting the model
      - Translational research (in converting theories to practice need to understand practice)

Example HIV education in South Africa

mukoma, fisher, ahmed et al.

- Intervention at the high school level. Provide education to:
  - Delay sexual debut / subsequent activity.
  - Increase correct condom use.
- You know the program was implemented.
- But you want to open the black box.
- How?

What's in a process evaluation 1

- Target participation
  - How much of target population was served?
    - Are other (unexpected) populations being served?
    - Are members of the target population aware of the program?
  - Is program filled to capacity?
    - Coverage: extent to which participation by the target population achieves the levels specified in program design
      - Undercoverage: WIC (can't include everyone in need)
      - Overcoverage: Sesame Street
    - Bias: Degree to which some subgroups participate in greater proportions than others
      - "Creaming or cream-skimming" (job training programs)
      - Self-selection (most motivated, aware, accessible)
  - Instead of prospective needs assessment = current NA
    - Admin records (who is in program, who drops out, who gets what services)
    - Survey clients (participants)
    - Community survey (only way to get at who is "out")

What's in a process evaluation 2

- Delivery of intervention
  - What is the program doing (activities)? (Service utilization)
    - Use program theory to decide which activities to focus on (which are most important)
      - Important to identify/collect data on
        - » Obstacles: what got in the way of delivering the services
        - » Facilitators: what helped the intervention get delivered
  - Non-program vs. Wrong program vs. Non-standardized
  - Specification of services and accessibility
  - Data strategies:
    - Review documents
    - Interviews (staff, participants, key informants)
    - Administrative data (program monitoring data, MIS)
      - Add items to existing data systems
    - Surveys (staff, participants)
    - Focus groups
      - good at identifying and reconciling different perspectives
      - Shared understandings
    - Observation
      - Follow a patient / subject (shadow)

Questions asked about services/activities

- Are participants receiving the proper amount, type, and quality of services?

- Are necessary program functions being performed adequately?

- Does the program coordinate effectively with the other programs with which it must interact?

- Are participants satisfied?

- Do participants engage in appropriate follow-up behavior after services have been delivered?

What's in a process evaluation 3

- What (organizational functions etc.) supports the program?
  - Program infrastructure
    - Governance
    - Staff (training, credentials etc)
    - Delivery system

Questions asked about organizational functions

- Are resources, facilities, and funding adequate to support important program functions?
- Are resources used effectively and efficiently?
- Is staffing sufficient (numbers and competence) for the functions that must be performed?
- Is the program well organized?
- Does program performance vary across sites?
- Is the program in compliance with applicable professional and legal standards?
- *Other questions?*

How to do process evaluation

- Describe/Document
  - Activities and those served and unserved
  - Outputs
- Identify Discrepancies
  - Compare program theory/logic model to actual
    - Every step/process
  - Fidelity
    - Reasoned vs accidental/inappropriate deviations
  - Standards
    - What is satisfactory?
      - Defensible standards or criteria
      - If don't exist: Program Theory is Critical
- Make Comparisons
  - Across sites
  - Model variations

Program monitoring

- ## Ongoing process evaluation
  - ## Repeated measurements over time
    - ### Program Process Monitoring
      - Systematic and continual documentation of key aspects of program performance that assesses whether the program is operating as intended or according to some appropriate standard
    - ### Program Outcome Monitoring
      - Continual measurement of intended outcomes of the program
        - » (often the social conditions it is intended to improve)

Example revisited

- Target population
  - Confirm teachers attended teacher training.
  - Confirm students were present during our treatment.
  - Try to confirm that teachers not in our program didn't know about it, and likewise students (how? By timing and location of training and classes).
- Delivery of intervention
  - Confirm that teachers grasped the key ideas with a "test" afterwards.
  - Confirm they were teaching lessons, and correctly, in class by spot checks.
    - Obstacles: teachers found some lessons hard to teach.
    - Facilitators: giving them material / props helped.
  - Confirm student attendance when classes taught by our own roll call.
  - Also review/ assist with prep and rollout of teaching materials and teacher training by talking to staff and teachers.
  - Ex post confirm that teachers believe what they were teaching; test students to confirm they have understood and believe the lesson.
  - For prizes: confirm students like the prizes; confirm they are available; and delivered as and when promised and that students know this

Example continued

- Encourage teachers continue with ideas from lesson even in other classes, but difficult to confirm.

- What else is displaced by our treatment? Was it important?

- Ex post debrief after pilot: did teachers/ student like the module and prizes. What worked/ didn't.

- What can be done to enhance teacher training and delivery of the modules?

Implementation evaluation

- More narrow definition of process evaluation
  - Rossi: Process=Implementation
  - Were services implemented?
    - vs. focus on the processes linking services to outcomes
- A broader definition of process evaluation
  - Whether services were delivered in such a way as to lead to the immediate outputs that then in turn will produce the desired outcomes

Evaluability assessment

- Systematic process that helps identify whether program evaluation is
  - justified
  - feasible
  - likely to provide useful information
- Shows whether
  - a program can be meaningfully evaluated
  - conducting the evaluation is likely to contribute to improved program performance and management. An evaluator needs to answer some important questions
- Before a process and/or outcome evaluation takes place

Example: YUVA Unstoppable

- A somewhat naïve and somewhat young researcher stumbles upon a seemingly revolutionary idea: arbitrage highly educated tech industry workers to coach failing schools physically next door.

- He decides to evaluate it and program agrees.

- RA is hired and spends a summer there.

- Prima facie program seems significant scale and impact.

- As summer evolves, new ideas are batted about for what the program can do.

- One month before the evaluation is to begin, program begins to worry about design (excluding some schools).

- Then when at is solved, one week before treatment program still doesn't yet know which schools will participate.

- Ex post, pretty clear it was not evaluable: too immature (treatment changing often), too unpredictable (seat of pants operation), too unprofessional (only volunteers, no paid employees), not that interested ultimately in being evaluated.

How the World Bank looks at evaluation

- Impact evaluations are part of a broader agenda of evidence-based policy making.

- Development programs and policies are typically designed to change outcomes, for example, to raise incomes, to improve learning, or to reduce illness. Whether or not these changes are actually achieved is a crucial public policy question but one that is not often examined.

- They provide a core set of tools that stakeholders can use to verify and improve the quality, efficiency, and effectiveness of interventions at various stages of implementation, or in other words, to focus on results.
  - Feedback loop for continuous improvement

World Bank view

- Monitoring is a <u>continuous</u> process that tracks what is happening within a program and uses the data collected to inform program implementation and day-to-day management and decisions = Process Evaluation
  - How well are programs/policies being implemented?
    - Especially: Alert to difficulties in implementing new policies
  - Descriptive studies
    - Challenges faced by state and local agencies when implementing accountability provisions of a new law
  - Crucial to understanding impact of policy changes but no substitute for "determining" impact.

Back to HIV/AIDS example from SA

- How to evaluate the process?
- Gather data:
  - Students from 12 schools + their parents.
  - Teachers: 25.

Results

Table II. Teachers' reports of number of lessons implemented at each school.

School	Number of students at the start of intervention	No. teachers involved in implementation[a]	No lessons implemented/ total	Comments
1	208	5	15/16[b]	Lesson 12 (condom use and demo) not implemented
2	340	3	16/16	Three classes had the intervention delivered by a teacher who did not attend any of the SATZ training
4	256	3	16/16*	Lesson 12 implemented by an outside facilitator (nurse)
7	153	1	16/16[c]	The same trained teacher delivered the intervention from beginning to end
8	410	3	15/16	Lesson 16 omitted as requested by programme designers
9	364	4	16/16	
11	254	5	15/16	Lesson 16 omitted as requested by programme designers
12	109	2	14/16[bc]	Lesson 12 (condom use & demo) not implemented; No lesson logs completed; no focus group discussions with students; The same trained teachers delivered the intervention from beginning to end
13	238	4	15/16	Lesson 16 omitted as requested by programme designers
18	208	1	16/16[c]	The same trained teacher delivered the intervention from beginning to end
22	313	2	16/16	2 classes had the intervention delivered by a teacher who did not attend any of the SATZ training
23	241	2	16/16[c]	The same trained teacher delivered the intervention from beginning to end
26	184	3	16/16	

Results

III. Teachers' mean scores on fidelity of lesson implementation: $n = 13$ lesson logs[a]. To what extent were the lesson/ac nented as planned?

n	Activity or lesson★				
	Mean (SD)[b]				
	1	2	3	4	5
Values clarification with regard to adolescent sexuality	3.5 (0.53)	3.4 (0.7)	2.6(0.9)	2.7(0.9)	2.4(1.2)	2
Self-esteem and sexual decision-making	3.7 (0.47)	3.4 (0.7)			
How our bodies function reproductively	3.4 (0.7)	3.3 (0.6)			
Dimensions of sexuality	2.7 (1.4)	2.1 (1.4)			
Boys don't cry! Girls are soft!	3.6 (0.5)	1.8 (1.8)			
Responsible decisions for sexual safety	3.5 (0.7)	2.2 (1.8)			
Promoting the sexual health of young people	3.5 (0.7)	2.6 (1.4)			
How do I handle this?	3.2 (0.8)	2.2 (1.7)			
Situations that carry the risk of sexual intercourse	3.8 (0.4)				
Coercion and violence in romantic relationships	3.2 (0.4)				
Not for me, not now!	3.5 (0.8)				
How to use condoms	3.0 (1.5)				
Negative consequences of sexual intercourse	3.2 (0.8)				
HIV and AIDS and the future	3.5 (0.8)				
Substance use and sexual decision-making	3.5 (0.8)				

Results

- Qualitative evidence from kids/ parents.
- What is missing here? Is it needed?

Process evaluation of the project on *Defining the Architecture and Management of a Global Subsidy for Antimalarial Drugs*

- Institute of Medicine report recommended the creation of a global-level subsidy for the new category of antimalarial drugs, artemisinin-based combination therapy (ACT).
- It recommended the establishment of a global fund that would purchase ACTs from manufacturers at a dollar price per dose and resell it at one-tenth of that price.
- The subsidized ACTs could be purchased by both the public and private sectors of all malaria-endemic countries.
- The subsidy would solve two critical problems at the same time:
    - it would enable widespread access to effective antimalarials (to save lives)
    - and would delay the emergence of resistance to artemisinin (to buy time).

Description of the global subsidy project

The Global Subsidy Project goal was to establish a viable plan for a global subsidy for effective combinations of antimalarial drugs. The project had three objectives:

1. To develop a detailed architecture and operational plan for a high-level global subsidy for effective antimalarials, including exit clauses to address situations in which the subsidy might no longer be needed or appropriate.
2. To build a coalition that has the critical mass to generate funding and political support so that the subsidy can become reality.
3. To address questions related to external efforts of the subsidy or external risk factors that could jeopardize the initiative.

Terms of reference for the evaluation

a. To assess the extent to which the project has met its stated
   goals and objectives, and assess the extent to which the
   project met any explicitly stated modification of those goals.
b. To measure these achievements against the most plausible
   counterfactual, i.e., what is the most likely scenario if the
   project had not been undertaken?
c. To describe the approach to the project, including the
   contents and processes, and assess these against best-in-
   class comparators in development assistance.

Evaluation design

Evaluation Design

Evaluation research questions	Indicators	Data sources
Research Question 1: Was a detailed architecture and operational plan for a high-level global subsidy for effective antimalarials developed? (related to project objective 1)		
1a: Was a detailed architecture and operational plan for a high-level global subsidy for effective antimalarials developed, including exit clauses?	• An architecture & operational plan was completed **(See evaluation document, p. 7)**	• Documents related to the AMFm Interviews with AMFm stakeholders
1b: Who developed the architecture and operational plan? What activities did these actors undertake, and how did these activities affect their ability to complete the architecture and operational plan?	• Number and scope of activities undertaken (including consultations and background papers commissioned) **(See evaluation document, p. 7)**	• Documents related to the AMFm Interviews with AMFm stakeholders
Research Question 2: Was a coalition of donors and political support built? (related to project objective 2)		
2a: Was a coalition of donors built who were i) willing to fund the subsidy, and ii) willing to work together on the subsidy?	• Presence of coalition of key donors in global health with the goal of working together on the subsidy **(See evaluation document, p. 7-8)** • Financial resources committed by donors for the subsidy **(See evaluation document, p. 8)**	• Documents related to AMFm • Media reports related to the AMFm Interviews with AMFm stakeholders
2b: What are the perceptions of the global subsidy for antimalarial drugs idea by key groups in global health? How do these perceptions match the project's view of the subsidy idea? Key groups include: i) donors; ii) NGOs in donor & recipient countries; iii) Ministries of Health in recipient countries; iv) international organizations; v) academics; vi) manufacturers, etc.	• Perceptions of key groups in global health about the subsidy idea **(See evaluation document, p. 8-9)** • Perceptions of project team about the subsidy idea **(See evaluation document, p. 8-9)**	• Documents related to AMFm • Media reports related to the AMFm Interviews with AMFm stakeholders
2c: Has the project managed these perceptions in terms of	• Number and scope of meetings and individual	•Documents related to AMFm

Evaluation results

- This process evaluation found that the Global Subsidy Project work program was essential to achieving the goal of establishing a viable, funded, global subsidy entity (the AMFm) at the GFATM. Specifically, the work program was based on a "Good Idea" and had policy champions and sufficient resources to propel the work forward. The actors conducting the work activities sought legitimacy in different organizations, demonstrated organizational learning, and took advantage of political opportunities in a changing global health context.

Sussman *et al.*: implementation

- Community-wide drug abuse prevention, involving schools, businesses, the media, and other units, may together provide the widest reach in helping to monitor drug use-related problems…

- This paper looks at program (Project Toward No Drug Abuse) that treats continuation high schools in CA as a community and implements such a broad-based program.

- 21 high schools randomly assigned to:
  - standard care,
  - classroom only (nine-lessons);
  - classroom + school as community (community speakers, peer meetings)

Sussman *et al.*: implementation

- Implementation evaluation: 5 hurdles:
  - Enlistment of school staff
  - Organization of associated student body
  - Establishment of student chair
  - Engagement in weekly meetings
  - Planning and completion of at least 6 events

- Process evaluation refers to "the perceived quality of a program on dimensions such as enjoyability… relevance…

PROCESS MEASURES

The weekly meeting form for ASB members contained nine items and fit on a single page. At the end of each ASB meeting, the student indicated the date and his or her school, and listed the activities that had been discussed during the meeting. The next eight items were closed-ended. These items were as follows:

1. Today's meeting was "very productive" to "actually counterproductive" (5-point scale).
2. I enjoyed today's meeting "very much" to "not at all" (4-point scale).
3. The staff person appeared to enjoy today's meeting "very much" to "not at all" (4-point scale).
4. Other students appeared to enjoy today's meeting "very much" to "not at all" (4-point scale).
5. Drug abuse was discussed today—"yes" or "no."
6. Drug abuse prevention (ways to keep nonusers from starting) was discussed today—"yes" or "no."
7. Drug abuse cessation (ways to help users to stop using) was discussed today—"yes" or "no."
8. Friends were brought to the meeting—"yes" or "no."
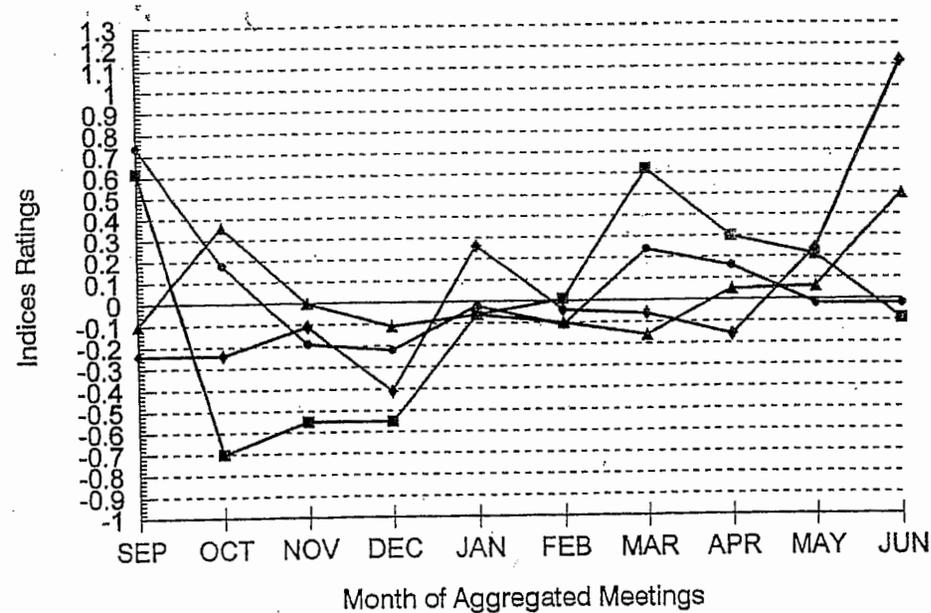
Sussman *et al.*: process, individual

Figure 1a: Standardized School-Aggregated Student and Staff Meeting Ratings, Plotted as a Function of Time (months)

NOTE:
- Student Productivity or Enjoyment; ▲ Student Drug Abuse Focus;
- Staff Productivity or Enjoyment; ◆ Staff Drug Abuse Focus.

Sussman et al: implementation, school

TABLE 3: Number of Reports of Activity Exposure by Condition

Type of Activity	Condition		
	Standard Care	Classroom Only	School as Community
Red Ribbon Week	11	9	10
Counseling groups	11	7	11
Recreation	6	2	18
Speakers	9	2	9
Drug education/ prevention/DARE	7	1	6
TND	—	3	8
Smoking cessation	1	2	6
Assemblies	4	3	—
Seminars	2	1	2
Job training	1	—	2
Fundraising	—	—	2
Staff training	1	—	—