

CONTEXTUALIZING CAUSAL THEORIES FOR PROCESS EVALUATION: WHAT MAKES OR BREAKS THE PROGRAM?

Michael J. Culbertson
University of Illinois

June 28, 2013

1

Workshop presented at the 2013 biennial meeting of the Society for Community Research and Action, Coral Gables, FL.

The title I didn't know I wanted:

**GREASE AND GRIME IN THE BLACK BOX:
LEARNING WHY PROGRAMS WORK BETTER
FOR SOME THAN OTHERS**

After I submitted the proposal for this workshop, I realized that I had a better title.

“Black box” is a reference to a common criticism of impact evaluation (evaluation that measures the size of a program’s effects), that it treats a program as a big black box—you just hit the button and get results (that the evaluation measures). Many evaluators want to “open up” the black box to take a look inside—to understand *how* the program achieves its outcomes, not just how big the outcomes are (often called process evaluation). We’re going to take things a step further and look at the contextual factors that facilitate or hinder the program’s functioning (what I’ll call “grease and grime”).

OUTLINE

- Motivation
- Mechanisms in Program Theory
- Case example: Community Knowledge Workers
- Hands-on exercise, part 1
- Grease and Grime
- Hands-on exercise, part 2
- Discussion

INTRODUCTIONS

- Find two people you don't know.
- Ask:
 - Name
 - Affiliation
 - Why they chose this workshop

About me: studied International Educational Development at Boston University, currently study educational research methods at the University of Illinois. Take a broad view education, includes schools, non-formal learning (e.g. afterschool programs), health education, agricultural education, etc. Career aim: help educational decision-makers understand and use information about their programs to make better decisions to improve educational access and quality.

MOTIVATION

Why doesn't a program work equally well for everyone?

When evaluating social programs, policy makers often ask the question “does the program work,” implying that the program either “works” or “doesn’t work” for everyone (or most people). But, just as people are not all the same, a given program isn’t going to work equally well for everyone. Evaluations that look at program impact or effectiveness many times examine only the average effect of the program, ignoring the question of who benefits most and who misses out. This workshop is all about that overlooked question: Why doesn’t the program work equally well for everyone?

Pair up with one other person and come up with a few reasons a program could work better for some than others. Try to think of a specific example in a program with which you are familiar. [5 min. Whole group share.]

What could you do if you had good information about why a particular program works better for some than others?

[Solicit whole-group responses.]

Now, we just heard a moment ago some hypothesized and anecdotal reasons why a program may work better for some than others. But, if we want to make good policy/programmatic decisions, we’ll need to gather good empirical evidence about exactly which potential influencing factors are most important for success in the program. Today, we’ll be using program theory as a foundation for conducting this kind of investigation.

PROGRAM THEORY

- Detailed description of HOW program activities lead to program outcomes/impact
- Causal chain / Mechanism
- Theory guides evaluation design

Program theory-based evaluation has been around for quite a while, but has received more attention in the last couple of decades. At the risk of oversimplifying, a program theory is basically a detailed description of HOW program activities (are thought to) lead to outcomes or impact. They often take the form of diagrams with events connected by arrows that are intended to convey the flow of events during and after a program, which you could call a “causal chain” (this event leads to this state change, which leads to this event, which leads to that outcome). I’m going to call this causal chain a *mechanism*. Program theory-based evaluation practice uses an elaboration of the program theory to guide evaluation design (what to look for, etc.).

MECHANISMS



My usage of *mechanism* is slightly different than others in program-theory based evaluation. Some authors [such as Weiss and Pawson/Tilley] say that the mechanism is what happens *after* the program activities have finished. For them, the mechanism consists of the intervening steps between the activities and the final outcome, and some [Pawson/Tilley] use the language of the program “firing” the mechanism, somewhat like a canon: The program activities unleash the mechanism as a single package that (assuming all goes well) travels along a set path to the destination (the outcome).

Instead, I prefer to think of a mechanism firing, not like a cannon, but like a neural pathway. The brain is made of up billions of neurons, each of which is connected to others in a vast network. When a set of connected neurons “fire,” each neuron passes an electrochemical impulse down the chain to the next. Similarly, the mechanism of a program theory describes the chain of events or state changes that make up the flow of program activities and the resulting outcomes. I prefer this image because it suggests that mechanisms are dynamic (not static), that they have multiple components, and that those components can be connected in complex ways (multiple pathways, systems dynamics).

One of the primary tasks in program-theory based evaluation is to test one or more potential mechanisms. That is, researchers would make careful observations to determine whether actual events follow the theorized causal links. But, our goal here isn’t to test a particular program theory. (Remember, we’re trying to understand why the program works better for some than others.) Instead, we’re going to take a slightly different tack and assume that a particular theory has been well enough established.

EXAMPLE: COMMUNITY KNOWLEDGE WORKER PROGRAM

- Rural model farmer/community members
- Participatory recruitment
- Ag info accessed by smartphone
- Liaison between neighbors and information
- Conduct surveys for development NGOs



Okay, let's get into an example using the Community Knowledge Worker Program, a community-based agricultural education program run by the Grameen Foundation in rural Uganda.

In traditional agricultural education (extension), a small number of highly trained extension agents attempt to serve a large number of farmers over a wide area. In places like Uganda, this often means that most farmers never or only rarely see an extension agent, since they simply have too much territory to cover (among other reasons). The Grameen Foundation Applab in Kampala decided to turn this model upside down: Instead of a few highly trained agents, Grameen recruits a large number of lightly trained individuals and equips them with access to detailed agricultural information.

The Community Knowledge Workers (CKWs) are model farmers/community members, chosen through a participatory, community-wide recruitment process. Once selected, CKWs receive a smartphone from Grameen, along with a three-day training program to teach them how to use the phone to access a database of agricultural information. The CKWs then serve as liaisons between their neighbors and the agricultural information. CKWs are "volunteer," but receive performance-based monetary incentives (monthly target: 48 searches, 15 farmer registrations). CKWs also help conduct surveys of their neighbors on behalf of other development organizations; this helps Grameen recoup costs and creates a two-way information flow between development organizations and the local community.



Page 2 of the handout summarizes some key features of the program (CIPP).

I became involved with the CKW program last year when the Grameen Foundation asked the MEAS consortium to conduct an impact evaluation of the program. MEAS (Modernizing Extension and Advisory Services) is a USAID-funded project to conduct, collect, and disseminate knowledge about agricultural extension around the world. Paul McNamara at the University of Illinois is the MEAS program director, and Dan McCole at Michigan State University serves as the PI on the CKW evaluation.

Since CKW is a completely new model of extension, we were interested in learning more about how the model works in practice than a strict impact evaluation by itself might provide, so we incorporated a program theory-based process evaluation into the design. The goal of this part of the study was to understand how CKW-farmer interactions take place (when, where, what, who talks, etc.) as well as to understand the contextual factors that influence those interactions and the farmers' uptake of agricultural information.

SAMPLE LOGIC MODEL

Inputs	Activities	Outputs
CKW Phone	CKWs meet with farmers	Farmers learn new farm info
Agricultural Database	CKWs search database	
Solar power system		Outcomes
		Improved farm productivity
		Impact
		Reduced poverty

10

One common way of representing a program theory is with a logic model (or log frame). For the CKW program, it might look something like this (I just dashed this off quickly for illustrative purposes).

Mechanisms are related to logic models, of course, since they're both ways of describing program theories. They also both have a "flow" to them. But, a mechanism (in my use of the term) is most likely going to be much more detailed than the logic model—describing each step in the process. Take for example the gap between "farmers learn new farm info" and "improved farm productivity"—isn't something missing here (farmer has to take action based on the information). Moreover, mechanisms allow for multiple pathways and complexity (feedback cycles).

CKW MECHANISM



What might the mechanism look like for a CKW/farmer interaction in the CKW program?

Note, this is only part of the program's entire mechanism. The farmer's action should lead to improved crop yields (etc.), which should lead to improve food security/reduced poverty (etc.). There are also precursor steps, such as CKW recruitment and training. But, since our research interest was about how the CKW/farmer interactions occur, I've only shown the part of the mechanism of primary interest.

This mechanism happens to be quite linear, but mechanisms could also have branches and multiple pathways, as well.

WHERE DO YOU GET THE THEORY?

- Program documentation
- Social science literature
- Program staff, administrators
- Program participants
- Logical thinking

So, you might be wondering where one comes up with a program theory [and the mechanism's accompanying grease and grime].

In the case of the CKW program, time and resource limitations unfortunately prevented a thorough model-building exercise. It would have been ideal to include program staff and participants, but program staff were busy, and I spent all my time in Uganda getting the main study going. So, the theory I presented here is mainly just my own reflections on the description of the program I received from program staff and talking with CKWs and farmers. Certainly not ideal, but a start and better than nothing.

EXERCISE: PROGRAM MECHANISM

- Read through program highlights
- Hypothesize a mechanism for the program
- Record any assumptions

Okay, let's try out what we've talked about so far. On pages 3-5 are brief descriptions of three programs. Pick one of the programs to work with, and in groups of 3-5, hypothesize a mechanism for the program. Of course, you don't know all of the details of these programs—don't get caught up with what you don't know. Simply make an assumption about the program for the sake of the exercise and move on. You can record the assumptions at the bottom of the CIPP sheet.

The goal here isn't to have a perfect mechanism—it's just to have a working draft for the sake of practice. In a "real-world" setting, you would want to do this part in collaboration with other program stakeholders (e.g. staff, participants, etc.).

TIPS FOR DRAFTING MECHANISMS

- Forward technique
 - What results from this? (Multiple outcomes?)
- Backward technique
 - What is necessary for this to occur?
- Missing link technique
 - What happens between two steps?
- Hint: Start with listed activities and outcomes

14

Remember that even though the CKW program mechanism we looked at was quite linear, your mechanism doesn't have to be linear. Use branches and multiple pathways as appropriate.

WHERE ARE WE?

- Why does a program work better for some?
- Program theory describes how program achieves outcomes
- Mechanisms / causal chain (gears in the box)

Okay, let's take a step back to see where we are. We started by asking the question "Why does a program work better for some than for others?" We then said that program theory could be a starting point for investigating this question, since program theory describes how it is that the program achieves its outcomes. One way of representing a program theory is with a mechanism (or causal chain). If an impact evaluation treats an intervention as a black box, the mechanism makes up the gears in the box, the inner workings.

Now, when you see a diagram of a program mechanism on paper, it gives the impression that everything runs smoothly from start to finish, much like the image of the cannon. But, we know that in real life, all kinds of contextual factors can disrupt the operation of the mechanism for a particular individual—after all, that's why we're in this workshop today. This is another place where I find thinking about the firing of a neural pathway more helpful than the firing of a cannon: You see, when a neuron (a link in the chain) receives an electrical impulse, it *isn't guaranteed* to pass along the impulse to the next neuron—whether the signal continues depends on contextual factors, such as the other signals the neuron is receiving from other neurons. Similarly, as people are moving through a program's mechanism, various individual contextual factors may push them along or hinder them from moving forward.

GREASE AND GRIME

- Individual contextual factors
- Regulate flow through the mechanism
- Grease: facilitating factors
- Grime: hindering factors
- Moderators

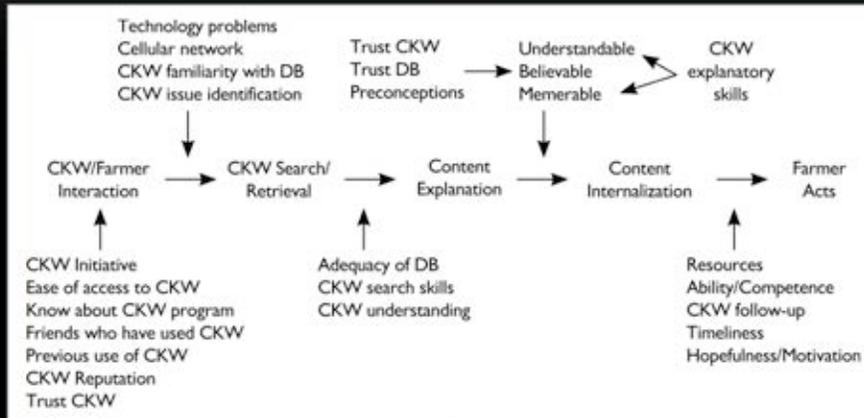
16

Switching back from the neural image to gears in the box, I'd like to introduce the concept of "grease and grime." These are individual attributes or contextual factors that regulate the flow of the mechanism. "Grease" includes things that facilitate the operation of the mechanism; and "grime" includes things that hinder its operation.

A quick side note: I imagine some people might get a bit squeamish with this talk of gears and such. After all, people aren't machines! In defense, I'd like to say that this is just figurative language to dress up the phrases "facilitating and hindering factors" in something a little more colorful. Of course people aren't machines—if they were, we probably wouldn't need to worry as much about the grease and grime, since they'd stick to the program mechanism as intended!

In any case, we can incorporate grease and grime into the model for our program theory as moderators in the mechanism, since they moderate the relationships between each step (that is, they make movement from step to step more or less likely to happen).

CKW GREASE AND GRIME



What kind of grease and grime is there in the CKW program?

[Solicit ideas from participants.]

Here's what I came up with....

Notice that grease and grime are often two sides of the same coin (good access to CKW is grease, bad access to CKW is grime).

Now we have a (considerable) list of things we think could account for why the CKW program works better for some than others. The next step is to collect empirical data about participants in the program to see which factors are most important in contributing to outcomes. This could be done in a number of different ways. In the case of the CKW evaluation, since we were piggy-backing on a study that involved extensive surveys, we decided to survey program participants (CKWs and participating farmers) about their experience with the program, using the list of grease and grime as a framework for generating the survey questionnaire.

TYPES OF GREASE AND GRIME

- What are some general kinds of grease and grime for:
 - Program participants
 - Program implementers
 - Program interactions/activities

18

We've seen some specific examples of grease and grime for the CKW program. Let's try to think of general categories of grease and grime that might be present in some form or other for a wide variety of programs. Get with a partner and come up with two or three general types of personal characteristics or contextual factors for participants, implementers, and interactions.

Page 6 of the handout lists some categories I came up with (suggestive, *not* exhaustive!).

EXERCISE: GREASE AND GRIME

- Identify potential grease and grime
 - What can the program respond to?
 - Hint: Focus on one link at a time
- Observation plan
- Implications for the program

19

Okay, let's give grease and grime a try with the mechanisms you worked on earlier.

Now, there are certainly hundreds of potential contextual factors that could influence a program. In order not to get overwhelmed and stay productive, focus on contextual factors that the program could respond to in some way.

Discussion

- What challenges did you note when identifying grease and grime for your program?
- How could you incorporate other stakeholders into the process?
- What challenges might you expect in observing/analyzing grease and grime for your sample program?
- How would you communicate results to stakeholders?
- What kind of evaluation environment would studying grease and grime be good for? Not terribly good for?
- How could studying grease and grime benefit a program you are familiar with?

CONTEXT IN PROGRAM THEORY-BASED EVALUATION

- Moderators
 - Generally sociodemographic (who, not why)
- Contribution Analysis (Mayne)
 - Context leveraged in attribution argument
- Realist Evaluation (Pawson and Tilley)
 - Small number of discrete contexts

21

Others in program theory-based evaluation have discussed the role of context. Grease and grime represent an elaboration of this discussion. Generally, context is considered in a small number of discrete units (e.g. sociodemographics). Grease and grime are intended to plunge the depths of psychological and interactional factors that may influence progress toward program outcomes. Moreover, grease and grime form a framework specifically for empirical investigation (as opposed to CA, where context is mainly an adjunct to the attribution argument).