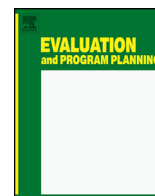




ELSEVIER

Contents lists available at ScienceDirect

Evaluation and Program Planning

journal homepage: www.elsevier.com/locate/evalprogplan

Causal inferences on the effectiveness of complex social programs: Navigating assumptions, sources of complexity and evaluation design challenges



Madhabi Chatterji

Teachers College, Columbia University, GDH 282 (P.O. Box 6), 525 W. 120th St., New York, NY 10027, United States

ARTICLE INFO

Article history:

Received 19 May 2016

Accepted 23 May 2016

Available online 18 June 2016

Keywords:

Impact evaluations

Experimental designs

Mixed methods

Causal inferences

Complex social programs

ABSTRACT

This paper explores avenues for navigating evaluation design challenges posed by *complex social programs* (CSPs) and their environments when conducting studies that call for generalizable, causal inferences on the intervention's effectiveness. A definition is provided of a CSP drawing on examples from different fields, and an evaluation case is analyzed in depth to derive seven (7) major sources of complexity that typify CSPs, threatening assumptions of textbook-recommended experimental designs for performing impact evaluations. Theoretically-supported, alternative methodological strategies are discussed to navigate assumptions and counter the design challenges posed by the complex configurations and ecology of CSPs. Specific recommendations include: sequential refinement of the evaluation design through systems thinking, systems-informed logic modeling; and use of extended term, mixed methods (ETMM) approaches with exploratory and confirmatory phases of the evaluation. In the proposed approach, logic models are refined through direct induction and interactions with stakeholders. To better guide assumption evaluation, question-framing, and selection of appropriate methodological strategies, a multiphase evaluation design is recommended.

© 2016 Elsevier Ltd. All rights reserved.

1. Purpose

In this methodological paper, I explore avenues for navigating assumptions, sources of complexity, and evaluation design challenges posed by a particular class of “evaluands” (after Scriven, 1991) that I call *complex social programs* (CSPs) here. A common example of a CSP would be a large scale, vaccination campaign to combat a rapidly-spreading disease in a rural community, or a comprehensive education reform initiative to improve student outcomes in an inner-city school district. My focus is on research design challenges and methodological obstacles faced by evaluators when conducting assessments of the impact of complex social entities such as the above, as encountered in their natural environments.

Impact evaluations, referred to more generally as effectiveness research in the social sciences, call for causal inferences regarding an intervention's effects based on outcomes manifested in program participants. Such studies seek to establish a causal link between the intervention and measured outcomes, coupled with a need to generalize the observed findings beyond the entities and

cases examined. By textbook traditions, the best research methods for making generalizable, causal inferences about an intervention's effects are experimental designs or variations thereof (Cresswell, 2008; Shadish, Cook, & Campbell, 2002). Of these, the “gold standard” of study designs is the Randomized Controlled Trial (RCT), a method that is highly valued in the traditions of evidence-based practice, stemming particularly from evidence-based medicine (Guyatt & Rennie, 2002).

The value for RCTs and a more general reliance on experimental or quasi-experimental designs for impact evaluations is both widespread and prevalent today. In a recently published Request for Applications from a prominent U.S. federal government agency, for example, we find the highest premium placed on evidence generated through RCT designs to help identify education programs, policies and practices that “work” (Institute for Education Sciences, 2015; see also What Works Clearinghouse at <http://ies.ed.gov/ncee/wwc/>).

CSPs are common to a number of fields, including education, business, social work, mental health, public health and medicine. Yet, the defining characteristics of CSPs – namely, the particular attributes that make such entities *complex* – are still nebulous in the methodological literature in evaluation. Correspondingly, there is limited organized knowledge on *how* theoretical assumptions of

E-mail address: mb1434@tc.columbia.edu (M. Chatterji).

textbook-recommended study designs (i.e., experimental designs), break down during impact evaluation attempts involving CSPs. In a similar vein, there is limited guidance on alternative but useful strategies and design tools to which researchers could turn, when performing impact evaluations of CSPs. Particularly, in instances where assumptions regarding how CSPs should work “in theory” (Chen, 1990), are in direct conflict with the fundamental assumptions underlying the textbook-recommended study designs, what recourses do evaluators and researchers have in generating a defensible body of evidence on what works or does not?

A central concern of this paper is to search for some cohesive answers to the above issues. I do so by building on my earlier work on similar themes (Chatterji, 2005, 2007, 2008, 2009; Chatterji, Kumanyika & Green, 2014), and by drawing on new literature on complexity theory from the areas of international evaluation, policy implementation, and the organizational sciences (Mason & Barnes, 2007; Matland, 1995; Nkwake, 2013; Patton, 2010).

1.1. Clarifying the problem

The task of the evaluator confronting a CSP is to generate a credible body of evidence to meet information needs of stakeholders (American Evaluation Association [AEA], 2011; Yarbrough, Caruthers, Shulha, & Hopson, 2010). A high priority for most program sponsors and funders – a primary stakeholder group – is for evidence of an intervention’s effectiveness as compared to the “treatment as usual” or the status quo. In many such cases, experimental designs are requested in explicit terms by stakeholders interested in making categorical decisions on program continuation (or discontinuation). Based on their theoretical advantages and utility in laboratory environments, there is an implicit faith placed in the scientific merits and rigor of experimental research. From many a sponsor’s perspective, it is only after an innovation is found to be sufficiently effective through rigorous experimentation, can funding be continued for making improvements, expansions, or scaling up.

However, experimental designs (particularly, RCTs) are extremely difficult to mount and sustain in real world settings for investigating any kind of social intervention, let alone CSPs. Numerous assumptions underlying experiments tend to become violated under field conditions, compromising the quality of evidence on both causality and generalizability. A few among these are that: complex programs are difficult to manipulate experimentally; initial randomization of subjects to treatment and control conditions is frequently compromised due to participant mobility or nonparticipation; environmental contamination of various types are a real threat, and issues of multiple causation—where more than one program influences the targeted outcomes in a given population—could potentially yield confounded, contradictory or un-interpretable effects. Sometimes early effects are in the reverse direction but completely misinterpreted by evaluators and researchers (e.g., where the control program may be more consistently implemented than the new one and shows better outcomes in early stages). Further, even when results favor the treatment, end users in evaluation contexts rarely find the limited information on “average effects” wholly satisfactory. To take actions on scaling up, future resource allocations or program continuity, decision-makers like to know how, when, why, and for whom a program works (or does not), and the costs and consequences of alternatives (Chatterji, 2005; Cook, 2002; Green & Glasgow, 2006; Green & Kreuter, 2005; West, Beisanz, & Pitts, 2000; Wolff, 2000).

Indeed, researchers and methodologists from a number of disciplines have documented issues affecting both internal and external validity of experimental results with social interventions

in general. Others have questioned the ethics of holding back necessary treatment from control group participants as a part of an experiment, when participants are from high need or high risk populations (Cook, 2002; Chatterji, 2005; Green & Glasgow, 2006; Green & Kreuter, 2005; West et al., 2000; Wolff, 2000).

Matters get further complicated when the evaluand is a CSP. In the typical case, CSPs are intended to serve as instruments of change and social betterment in large organizations or communities. Given this, numerous added threats to the *ecological validity* of findings arise if one relies only on experimental designs. As the forthcoming examples and case study in this paper will elaborate, CSPs rarely “follow a linear pathway” to the desired ends set by stakeholders in the settings where they are found (Patton, 2010, p. 5). Indeed, the flows of activity surrounding a CSP are often non-linear; interactions among actors are dynamic; big or small incidents in their larger environments can lead to big momentum shifts in realized outcomes; and both expected and unexpected results might ensue, catching both evaluators and stakeholders off-guard. Resistance, disorderliness and turbulence are natural responses to change in any social organization, and at different points in time, could lead to varying degrees of chaos, ambiguity, conflict and uncertainty in activities as well as outcomes (Matland, 1995). Paradoxically, the methodological issues arise *because* the complexities inherent in the very definition, purposes and larger ecology of CSPs are in direct conflict with the basic assumptions underlying experimental designs.

Given the above reality, a practical first step for evaluation researchers is a “situation recognition” exercise to distinguish among “Chaotic”, “Complex”, “Complicated” and “Simple” evaluation situations (categories derived from Patton, 2010, p. 80). To make optimal methodological and design choices when conducting impact evaluations of CSPs, evaluators must become comfortable in dealing with both complexities and assumptions in these evaluation environments (Nkwake, 2013; Nkwake & Morrow, editorial to this volume). Further, they must develop capacities in making effective methodological decisions that add rigor to the impact studies while meeting stakeholder needs for information.

This paper is intended as a logical extension to the lines of thinking initiated by Patton (2010), Nkwake (2013) and authors in this volume, as applied to impact evaluations of CSPs. It attempts to address the earlier deficits noted in the evaluation literature by searching for methods and tools to improve the rigor and quality of impact evidence of CSPs. My objectives are as follows:

- (1) Develop a common definition of a Complex Social Program (CSP) based on an analysis of field-based examples and cases from different disciplines—specifically, from public health, education, and mental health areas;
- (2) From an evaluation case analysis, derive (and categorize) the major *sources of complexity* that typify CSPs, identifying the corresponding design challenges and specific threats to generating sound evidence of impact; and
- (3) Propose and select the *best-suited* tools, concepts, and methodologies from a relevant literature review for collecting impact evidence on CSPs.

1.2. Method

I begin with a working definition of a CSP by identifying their distinctive attributes from three applied evaluation examples. I then present the assumptions underlying the standard experimental design, highlighting specific design-related assumptions that conflict with typical configurations of the CSPs described. Next, I analyze in detail an evaluation case involving a CSP, namely, a study of a global health program for migrant populations. From

Download English Version:

<https://daneshyari.com/en/article/6792426>

Download Persian Version:

<https://daneshyari.com/article/6792426>

[Daneshyari.com](https://daneshyari.com)