Evaluation and Program Planning 52 (2015) 118–125

Contents lists available at ScienceDirect

Evaluation and Program Planning

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

The power of the context map: Designing realistic outcome evaluation strategies and other unanticipated benefits

Ralph Renger [b], Jirina Foltysova [a], Karin L. Becker [a,*], Eric Souvannasacd [a]

[a] Masters of Public Health, University of North Dakota, Grand Forks, ND, 58202, USA
[b] School of Medicine and Health Sciences, University of North Dakota, Grand Forks, ND, 58202, USA

ARTICLE INFO

Article history:
Received 13 January 2015
Received in revised form 15 April 2015
Accepted 23 April 2015
Available online 7 May 2015

Keywords:
Context map
ATM
Program theory
Theory driven evaluation
Multiple activities
Outcome evaluation

ABSTRACT

Developing a feasible evaluation plan is challenging when multiple activities, often sponsored by multiple agencies, work together toward a common goal. Often, resources are limited and not every agency's interest can be represented in the final evaluation plan. The article illustrates how the Antecedent Target Measurement (ATM) approach to logic modeling was adapted to meet this challenge. The key adaptation is the context map generated in the first step of the ATM approach. The context map makes visually explicit many of the underlying conditions contributing to a problem as possible. The article also shares how a prioritization matrix can assist the evaluator in filtering through the context map to prioritize the outcomes to be included in the final evaluation plan as well as creating realistic outcomes. This transparent prioritization process can be especially helpful in managing evaluation expectations of multiple agencies with competing interests. Additional strategic planning benefits of the context map include pinpointing redundancies caused by overlapping collaborative efforts, identifying gaps in coverage, and assisting the coordination of multiple stakeholders.

Evaluators are frequently challenged with developing a feasible and realistic outcome evaluation (Mark, Henry, & Julnes, 2000). There are different evaluation approaches capable of assisting in meeting this challenge based in use, methods, and values (Alkin, 2012; Chen, 1990; Patton, 2008). The focus of this article is on adapting the Antecedent Target Measurement (ATM) approach (Renger & Titcomb, 2002), a methodology grounded in theory driven evaluation, to define the outcomes most likely to demonstrate impact by program activities.

Central to theory driven evaluation is program theory (PT). A PT makes the underlying assumptions of the program explicit (Blamey & Mackenzie, 2007; Donaldson, 2007; Renger & Titcomb, 2002). Program assumptions can be expressed as mechanisms of change, antecedent conditions, risk factors, contributing factors, and so forth (Chen, 1990; Leeuw, 2003; Renger, Bartel, & Foltysova, 2013; Weiss, 1997).[1] More specifically a PT identifies the underlying conditions of a problem being targeted by the program. This is important to the arguments and method presented below because the PT isolates the subset of underlying conditions being targeted by the program from all the underlying conditions defining the context in which a program operates.

To develop a PT, it is best, albeit not always possible, to begin by defining the context in which program activities operate (Renger, 2011). The context is defined as the conditions (e.g., dislike of exercise, sedentary lifestyle, etc.) underlying a problem (e.g., obesity). Once underlying conditions defining the context are made explicit, then activities can be meaningfully aligned to them (Renger & Titcomb, 2002). The underlying conditions targeted by the activities then define the immediate and intermediate outcomes of interest to be evaluated (Renger & Titcomb, 2002). The activities and outcomes together define the PT (Weiss, 1997).

In the authors' experience it is common for a program to target several, but not all, of the underlying conditions defining the context. Fig. 1 depicts a hypothetical PT of a program attempting to

* Correspondence to: Center for Rural Health, School of Medicine and Health Sciences, University of North Dakota, 501N. Columbia Road Stop 903, Grand Forks, ND 58202-9037, USA. Tel.: +1 701 777 4499.
E-mail addresses: ralph.renger@med.und.edu (R. Renger), Jirina.foltysova@med.und.edu (J. Foltysova), Karin.becker@email.und.edu (K.L. Becker), Eric.souvannasacd@med.und.edu (E. Souvannasacd).

http://dx.doi.org/10.1016/j.evalprogplan.2015.04.003
0149-7189/

[1] Hereinafter, for the purpose of consistency the term underlying conditions is used as an all-encompassing term.

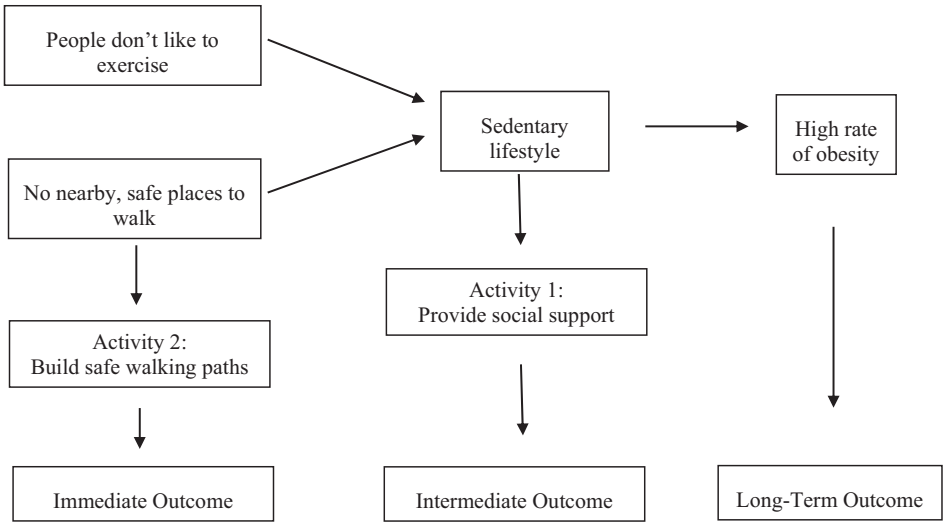

Fig. 1. A hypothetical program theory.

effect change in obesity via physical activity by providing safe walking paths and social support.

Most of the programs the lead author evaluated over the last two decades mirror the example in Fig. 1. There are, however, numerous other underlying conditions affecting physical activity not being targeted and/or considered, such as diet, genetics, and so forth. A narrow program and evaluation scope is often necessary because of restrictions to funding and program length (Bamberger, Rugh, & Mabry, 2006; Renger, 2011; Schalock & Thornton, 1988). Under such circumstances it is reasonable to expect changes in immediate outcomes because they are being directly targeted by the program. However, as one moves across the continuum of outcomes from immediate, intermediate, to long-term, the likelihood of demonstrating change becomes less likely. This is because a program operates in a broader context not depicted in the PT (Huntington & Renger, 2003; Morell, 2005, Morell, 2010) and there are many other underlying conditions not targeted by a program affecting the likelihood of demonstrating change in outcomes, but over which the program has no control (Huntington & Renger, 2003).

Some agencies are acutely aware of the broader context and the numerous underlying conditions affecting the likelihood of program activities having their intended impact. Therefore, they engage in a more ambitious effort to target additional underlying conditions by (a) expanding the reach of a single activity, and/or (b) incorporating numerous activities to target additional conditions. Many smoking cessation programs are good examples of where multiple interventions are used to address numerous underlying physiological, psychological, and social conditions (American Lung Association, 2014; Jefferson University Hospitals, 2014; Kansas Department of Health & Environment, 2014; Legacy, 2014; North Dakota Department of Health, 2013; Respiratory Health Association, 2014).

One important factor affecting the degree to which additional underlying conditions are targeted is resources. It is possible that a single agency is able to secure the funds needed for a broader program scope, but most often agencies must partner with other service providers to leverage the resources needed to increase programming breadth and reach. When this occurs the evaluator is confronted with additional challenges. First, the evaluation budget is rarely sufficient to evaluate all of the targeted underlying conditions. Second, if the necessary evaluation budget were available, then the expanded scope of the evaluation plan poses significant ethical and feasibility concerns. For example, often more staff time is needed to assist in carrying out an evaluation plan with a larger scope. This then reduces the time staff has to provide services (Renger, 2014). Third, although each agency in a multi-agency collaborative has an interest in contributing to the whole, their primary interest is in evaluating the outcomes associated with underlying conditions targeted by the investments they are making. This creates significant pressure on the evaluator to engage all agencies in a fair process and maintain participant motivation throughout the evaluation even if an individual agency's outcomes of interest are not represented in the final evaluation plan.

In the authors' two decades of experience, these challenges presented themselves in evaluating the Housing and Urban Development Housing Opportunity for People Everywhere program (Renger, Passons, & Cimetta, 2003), the National Science Foundation funded Partnership for International and Research and Education (Renger & Foltysova, 2012), and the Centers for Disease Control and Prevention and administered through the Arizona Department of Health Services (Renger, Kidd, & Jansen, 2006).

The theory driven evaluation literature was reviewed for solutions to assist in defining and evaluating outcomes when multiple activities and/or agencies work together toward a common goal. A common and useful theory driven evaluation method is the logic model. The logic model summarizes the "logical" process of linking underlying programmatic assumptions, activities, and outcome measures (Donaldson & Lipsey, 2006; Renger et al., 2013; Renger & Titcomb, 2002). There are different types of logic models ranging from table-format style (W.K. Kellogg Foundation, 2010) to visual maps (Renger & Titcomb, 2002; Rosas, 2005).

The table-format works well for showing the logical connections between a small number of activities and targeted underlying conditions. However, the table-format quickly becomes unwieldy as the number of underlying conditions and activities to be evaluated increases (Funnell & Rogers, 2011).

Further, as Morell (2014) noted, one unintended consequence of the table-format logic model is, it does not show the relationship between the program and its environment. That is, it only depicts the subset of contextual conditions being targeted by the program. It does not show the broader context of the other underlying conditions contributing to the problem, but are not being targeted.

Another pitfall of the table-format logic model is "retrofitting". In retrofitting the activity is predetermined and the programmatic assumptions are made to "fit" the activity. The result is a "tight"