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).

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

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

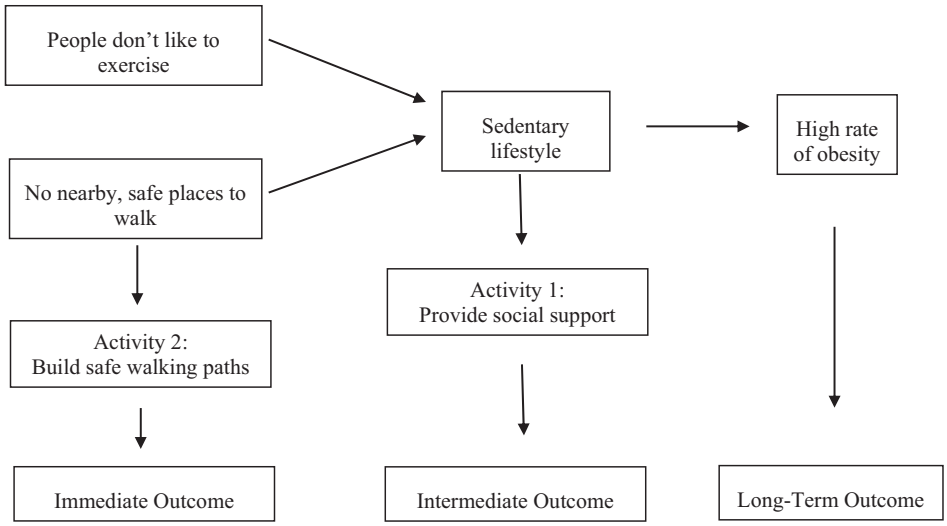

Fig. 1. A hypothetical program theory.

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"

logic model, one where programmatic assumptions, activities, and outcomes are perfectly aligned. This is then showcased as the justification for the program (Hummelbrunner, 2010). Since the purpose of providing a logic model is justification, retrofitting does not ensure the most salient underlying conditions are being targeted.

A method was needed for incorporating context to develop a realistic outcome evaluation in cases where multiple activities work toward a common goal. The solution to meet these challenges was to adapt the ATM approach for logic modeling (Renger & Titcomb, 2002). The ATM approach continues to be a widely used theory driven evaluation method in program evaluation. According to Google Scholar® the theoretical and practical application article by Renger and Titcomb (2002) has been cited 123 times by other evaluation manuscripts.

Some underlying assumptions relating to developing an outcome evaluation for programs with multiple activities are first shared. Then, a case example illustrates how the ATM approach was adapted to develop a realistic and feasible outcome evaluation plan for a multi-agency obesity collaborative. A total of 12 agencies formed the collaborative. To facilitate understanding and to protect anonymity the case illustration uses just three hypothetical agencies each responsible for implementing a different activity.

Method assumptions

Assumption 1: A realistic outcome evaluation begins by first understanding the context in which a program operates.

As stated above Morell (2014) notes one unintended consequence of the table-format logic model is it neglects context. Understanding context is essential in defining immediate, intermediate, and long-term outcomes and more importantly to establishing the likelihood that each will change.

Assumption 2: No program can ever be designed to target all underlying conditions of a problem.

It is typical to identify between 50 and 100 underlying conditions contributing to a social problem (Foltysova, 2013; Renger, 2011). Most programs, due to resource limitations (e.g., budget, expertise, etc.), can only attempt to exert control over a small subset of these underlying conditions (Huntington & Renger, 2003).

Assumption 3: When there are multiple activities they are designed to work together in a coordinated way toward a common goal.

The purpose of bringing multiple activities to bear on a problem is to (a) increase the likelihood of changing a single underlying condition by increasing the intensity with which it is targeted (i.e., repeatedly and in different ways) or, (b) target a greater number of underlying conditions thereby improving the likelihood of seeing change in the goal, or (c) both a and b.

Adapting the ATM approach: a case illustration

Step 1: Generate an understanding of the context

The first step of the ATM approach engages subject matter experts in root cause analysis (RCA) to identify and visually depict relationships between the problem and Antecedent, or underlying conditions. RCA is a straightforward method focusing on digging deeper by asking a series of "why?" questions (Coşkun, Akande, & Renger, 2012). Clarity on the problem statement is essential to the success of the process (Foltysova, 2013; Renger, 2011; Weiss, 1997). Limited or lack of consensus about the problem statement can stall or derail the process leading to numerous antecedent conditions being identified that are unrelated to the problem statement.

Possible candidates to be considered as subject matter experts are professionals holding leadership positions, program staff members, expert scholars, and business executives. Selecting appropriate subject matter experts can be difficult. Patton (2008) recommends relying on leadership to identify key subject matter experts. In our RCA interview process, we document years of experience, title, and training for each subject matter expert.

The goal of RCA interviews is to visually capture the context by identifying as many of the underlying conditions that contribute to the problem (and the relationships between them). Results of individual RCA interviews are then integrated into a single, summary image referred to as the context map. As noted earlier, it is common for the final, summary, context map to depict upwards of 50–100 underlying conditions (see Fig. 2).

During RCA interviews it is natural for a subject matter expert to focus on making explicit the underlying conditions being targeted by his/her respective agency. However, for the methodology to be successful subject matter experts must be encouraged to identify as many conditions underlying the problem including those extending beyond the scope of the specific activities in which they are invested.

Fig. 2 depicts a context map of underlying conditions related to the problem of obesity identified by subject matter experts. As a reminder, the context map is a necessary precursor to the PT, but is itself not the PT.

The RCA process is not perfect; sometimes underlying conditions are missing or relationships between them are not accurately depicted (Foltysova, 2013). Therefore, it is important to validate the PT using member checks (Lincoln & Guba, 1985; Renger & Bourdeau, 2004). A member check is an interview summary highlighting the subject matter experts' major points. The member check is often in narrative form. Supplementing the narrative with the context map significantly improves the response rate and feedback quality of the member check (Foltysova, 2013; Renger & Titcomb, 2002). It is best to send the member check to the subject matter experts immediately following the interview before the recollection of the interview and process fades. The member check is advantageous because often subject matter experts recall other important issues after the interview. The member check provides subject matter experts an opportunity to add contextual factors or correct the understanding of the relationship between them. Member checking is critical for buy-in and ensures the final product is deemed credible and used (Patton, 2008).

Step 2: Align activities to the context map

It is likely many of the underlying conditions in the context map are beyond the direct and immediate control of the program to impact because of resource, expertise, and time limitations. Therefore, Renger and Titcomb (2002) suggest engaging 5–10 key agency level decision-makers in a prioritization process. The goal is to narrow down the myriad of underlying conditions to a more manageable subset. For example those within the agency's mission, budget, and which are likely to show change in the funding cycle.

Based on the authors' experience the prioritization process does not work well when working with multiple agencies with different missions, goals and priorities. It is difficult to achieve consensus when there are so many competing interests. Therefore, another way to prioritize was sought. The problem was likened to triangulation, a common practice in qualitative methods to assess credibility and trustworthiness (Padgett, 2008).

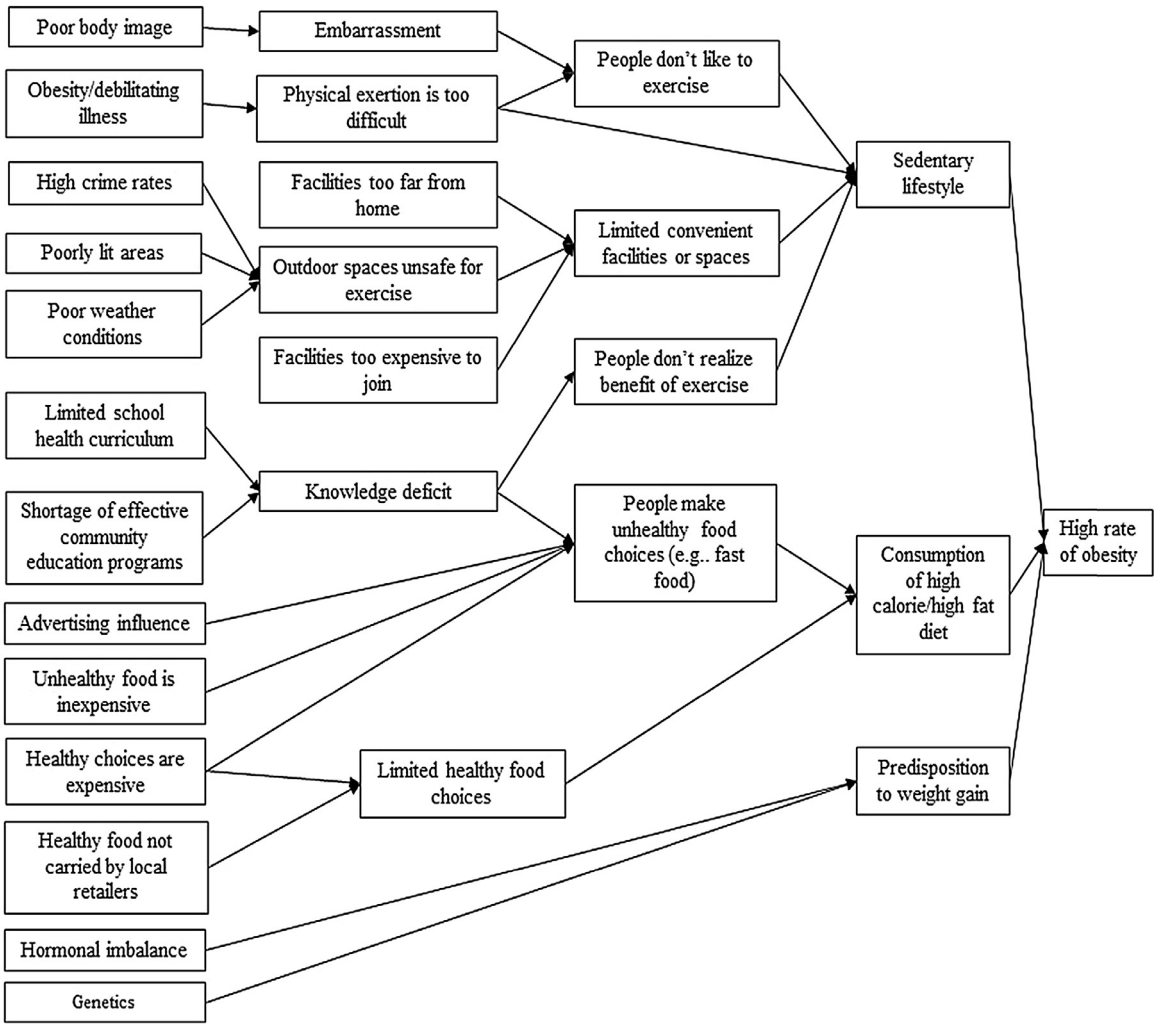

Fig. 2. Developing the context map.

As one of six strategies for rigor as outlined by Padgett (2008), triangulation involves employing two or three research methods tasked with the same question; if two or three methods produce similar results then the results are validated. Fielding and Fielding (1986) state, "Triangulation puts the researcher in the frame of mind to regard his or her own material critically, to test it, to identify its weaknesses, to identify where to test further doing something different" (p. 24).

This process of triangulation fuses anecdotal evidence alongside quantitative data to provide a "reasonably viable portrait" of findings and is consistent with convergent analysis techniques and multi-method approaches (Nguyen, Attkisson, & Bottino, 1983, p. 104). Triangulation also hinders relying exclusively on a single data collection, reduces bias inherit in a particular data set and discourages eliminating a data set for ease of analysis (Anfara, Brown, & Mangion, 2002). From the practice of triangulation, evaluators can assess findings with confidence. It is an extra tool to assess whether resources have been invested in the right places, ensures greater accuracy in identifying alternative explanations emerging from findings, and informs the basis of valid and reliable multiple case study results (Stake, 2006).

In the current context triangulation occurs when multiple activities are reviewed for areas of overlap and for patterns of agreement as well as areas of division. To determine the degree to which underlying conditions are targeted, or coverage, each activity must first be aligned to the context map. To do this it is suggested to begin by examining source documentation (e.g., a lesson plan for an activity) to determine which underlying condition(s) an activity is designed to target (Renger, 2011). Staff overseeing the activities should also be consulted, presented the context map and asked to identify the underlying conditions targeted by each activity. The targeted underlying conditions should then be highlighted in the context map (Fig. 3).[2]

A cautionary note: on rare occasion agency decision makers will overstate the number of underlying conditions being targeted and on a few occasions some claim their program activities "targets everything". Such comments are a good indicator the process is perceived as threatening or misunderstood. Agency decision makers can become concerned their activities will not be implemented if the underlying conditions targeted by their program's activity are not selected to be evaluated. Of course, this is not true. Therefore, reinforcing the notion that aligning activities to the context map is intended to direct finite evaluation resources, and not dictate operations, is essential.

Step 3: Define the immediate, intermediate, and long-term outcomes

Fig. 3 shows the underlying condition being targeted by three activities. The shaded boxes in Fig. 3 signify underlying conditions

[2] Color coding of activities is preferred. Journal formatting restrictions do not permit the use of color.

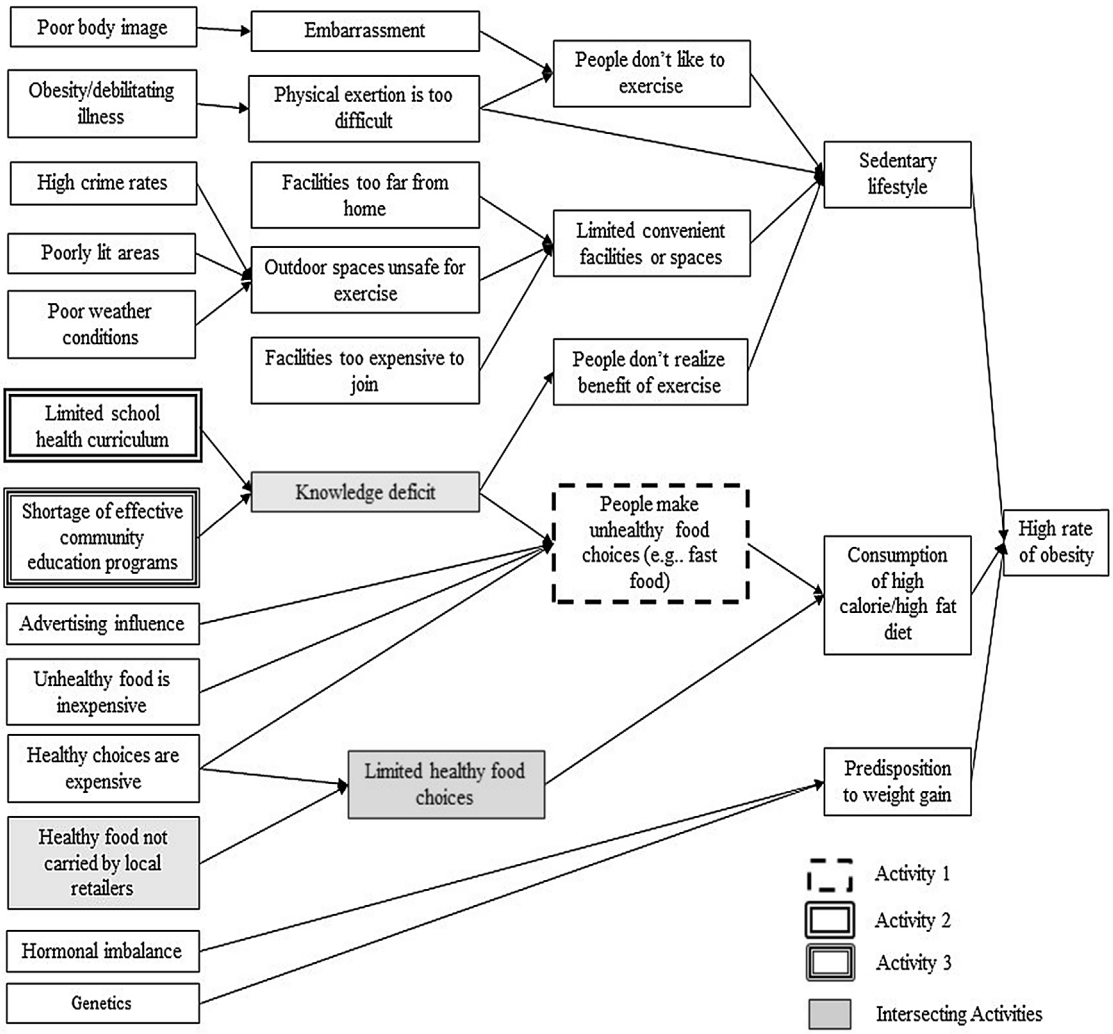

Fig. 3. Using the context map to prioritize.

targeted by more than one activity. The targeted underlying conditions then become the potential immediate, intermediate, and long-term outcomes to be evaluated. The matrix shown in Table 1 is designed to assist in prioritizing potential outcomes from the context map.

Defining the scope of the outcome evaluation depends on the simultaneous consideration of the degree of (a) coverage and (b) control the program potentially can exert over underlying conditions. If an underlying condition is not targeted by any activities, then clearly it should not be considered in the outcome evaluation. If an underlying condition is targeted by at least one activity, then it could potentially be included in the outcome evaluation. When an underlying condition is targeted by multiple activities it should become a priority for consideration in the outcome evaluation because considerable programmatic resources are being invested in attempting to impact it. When multiple underlying conditions are targeted by multiple activities, then priority should be given to evaluating the underlying condition beginning a thread (i.e., is furthest to the left in the context map). Underlying conditions beginning a thread are root causes. There is a greater likelihood of showing impact in a root cause because (a) it is being directly targeted by program activities, and (b) by definition there will be no competing underlying conditions attenuating the ability to demonstrate impact.

Table 1
Matrix for prioritizing antecedent conditions for outcome evaluation.

	High coverage (Two or more targeted underlying conditions)	**Low coverage (one activity targeting an underlying condition)**	**No coverage**
High (immediate) control	**High priority** e.g., Healthy food not carried by local retailer	**Medium priority** e.g., Limited school curriculum	**Do not evaluate**
Low (intermediate) control	**Medium–high priority** e.g., Limited healthy food choices	**Low priority** e.g., People make unhealthy food choices	**Do not evaluate**
No control	**Do not evaluate**	**Do not evaluate**	**Do not evaluate**

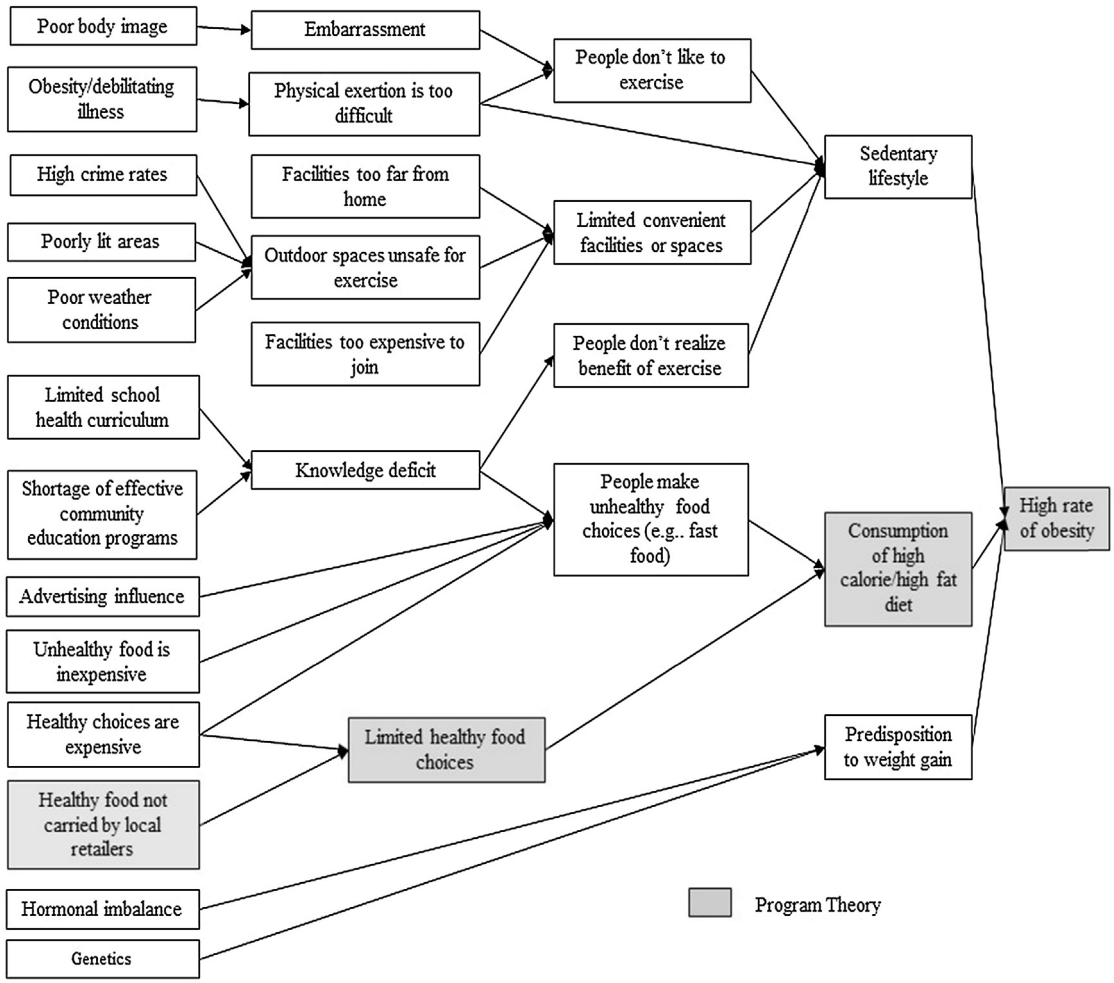

Fig. 4. The PT in context.

Step 4: Define the PT

Fig. 4 shows the hypothetical results of the prioritization process. Only the underlying conditions surviving the prioritization process are included in the final PT. The final PT is reviewed jointly one last time by the agency decision makers. It is important to again remind agency decision makers the prioritization process is intended to better direct finite evaluation resources; it is not a program moratorium. This is especially important when multiple agencies are represented. Some agencies mistakenly conclude their program is of less importance if they do not see underlying conditions targeted by their program(s) represented in the final PT.

The prioritized outcomes are of course open to negotiation with the agency decision makers. If agency decision makers are adamant about the need to include certain outcomes, then these can be added to the evaluation plan. However, it is important the context map continues to be the basis for managing expectations regarding the likelihood of incorporating user added outcomes.

It is also important to point out the PT includes the entire thread of outcomes from immediate to intermediate to long-term, even though the intermediate and long-term outcomes were not directly targeted by any one activity. This is because it is necessary for a PT to depict the logical connection between the activity and the long-term outcome. However, from Fig. 4 it can be seen the power of using this approach is in depicting the PT in context. This makes it easier to (a) understand the logical relationship between activities and the ultimate program goal and, (b) develop realistic expectations regarding the likelihood certain outcomes can be impacted by the program activities. As one moves from immediate to intermediate to long-term outcomes the likelihood of demonstrating impact decreases because there are more underlying conditions over which the program(s) have no control.

Discussion

Understanding what outcomes to evaluate can be challenging for evaluators. The problem is made even more complex when multiple agencies are involved, each with their own stake in the evaluation, and operating with limited evaluation budgets. Under such conditions evaluators need a systematic and fair approach to identify the outcomes to be included in the evaluation plan.

The logic model is a useful tool, but has limitations that are magnified when attempting to evaluate multiple activities. The most notable problem is the table-format logic model does not depict the context in which a program is operating, thus the resulting PT may be oversimplified and lead to unrealistic expectations regarding expected changes in intermediate and long-term outcomes.

The adapted ATM approach shared here is appreciated, understood, and deemed fair by stakeholders. This conclusion is based on the fact only two of the over 45 agencies involved in the several collaborations dropped out because the final evaluation plan did not include underlying conditions they targeted. One reasonable explanation for the high retention rate is program staff

and agency decision makers (a) are the subject matter experts used to generate and validate the context map, (b) provide the source documentation, and (c) validate the activity mapping exercise.

The ATM approach and resultant context map are grounded in theory driven evaluation. The context map is intended to help understand those outcomes over which the program has direct and immediate control to change. Displaying the programmatic assumptions (i.e., mechanisms of change) in context helps clients establish realistic expectations regarding changes in immediate and intermediate outcomes. That is, as one moves from immediate to intermediate to long term outcomes the likelihood of observing change based on the program alone diminishes; there is an increasing number of factors impacting long-term outcomes over which the program has no control.

In the interest of transparency the primary purpose of developing the context map was simply to manage and establish realistic expectations for change in outcomes. However, the utility of the context map can be augmented by considering premises of another theory driven evaluation approach: realistic evaluation. Realist evaluation notes the importance of context in interpreting outcomes and postulates contextual factors may differentially affect the mechanisms of change (i.e., the program theory) which in turn differentially affect the outcomes (Pawson & Tilley, 1997). In this way the context map from a realistic evaluation perspective provides the evaluator a foundation that could be especially helpful in directing further inquiry should certain outcomes not be achieved. For example, to provide specific recommendations for substantive program changes addressing specific contextual factors that impact the ability to demonstrate positive outcomes. The process described here only assists the evaluator in prioritizing the outcomes to evaluate. The evaluator still must determine how to evaluate the outcomes.

There are several unanticipated benefits of using the adapted ATM approach. The first is the ability of the context map to facilitate strategic planning. It is rare to see Assumption 3 met; in practice multiple activities or multiple partners seldom work together in a coordinated way. Stakeholder partners working collaboratively are diligent; they do things right, but often are not doing the right things. The context map allows individual activities to be examined and tweaked as necessary thereby avoiding activity traps (Renger & Titcomb, 2002). When activities are aligned to the context map gaps and redundancies in underlying conditions become clear. In the case of gaps, the context map can be used to identify the correct partner to bring into a collaborative. In the case of redundancies, the decision can be made whether the multiple activities are needed to bring about change (i.e., increased intensity) or whether some activities could be dropped or redirected to target other antecedent conditions.

Lessons learned

One criticism levied against outcome evaluations is the failure of methods, such as logic models, to capture the context in which the program operates. Consequently, the expectations regarding the likelihood of observing change in intermediate and long-term outcomes may be unrealistic. One reason for not defining the context is the belief it is time consuming and costly. While there is an upfront cost for conducting RCA interviews with subject matter experts to define the context, it is relatively fast and inexpensive. Often 10 to 12 interviews of less than 1 h are needed to identify a host of underlying conditions impacting the program's likelihood of success. When validation checks and synthesis of subject matter interviews are factored in an efficient and competent evaluator can define the context using the adaption of the ATM method described here in less than 40 h (Renger & Hurley, 2006). These estimates are not contingent on the number of participating agencies; the determining factor is based on how many interviews are needed to reach a saturation point to understand the context. As long as there is a shared understanding of the problem across agencies these are reliable estimates.

There are three primary benefits of the context map to evaluators. First, by aligning activities to the context map it can guide the evaluator to allocate finite resources by identifying outcomes most likely to change. Second, engaging stakeholders in defining the context map is an effective strategy for building buy-in and thus improving the utility of the evaluation results (Patton, 2008). Third, although it is impossible to identify all the contextual conditions in which a program operates the process described here is more than sufficient to demonstrate there are numerous conditions over which the program has no immediate or direct control to change. This helps build realistic expectations regarding the outcome evaluation.

The context map also has benefits extending beyond evaluation, assisting stakeholders in strategic planning. The context map is especially useful in coordinating multiple activities and multiple agencies. As with anything in life, you are likely to get out only as much as you put in. We encourage evaluators to take the time to define the program context as the many benefits for evaluators and stakeholders offset the initial upfront investment.

We also caution evaluators not to assume overlapping activities work together in a coordinated way. Evaluators can assist clients by ensuring consideration is also given to implementation theory (Renger et al., 2013; Weiss, 1997). A sound implementation theory can serve to bind activities around a common theme. An explicitly stated implementation theory should also then be evaluated to ensure the synergistic relationship among multiple activities is achieved.

Conclusion

The concept of a context map is not new and there are variations found within evaluation (Rosas, 2005). One limitation of using the adapted ATM approach is the one-directionality of the context map. This occurs as a result of repeatedly asking why when using the RCA as the interview premise. Of course, in reality many relationships are bi-directional, multi-directional or nonlinear. Thus use of other methods like causal loop diagrams should be considered in augmenting the ATM approach to more accurately capture the context.

The unique contribution of the approach presented here is it builds on a proven theory driven evaluation approach to logic modeling to place the PT in theory context. Placing the PT in context provides a visual tool for evaluators far more powerful than the table-format logic model. Thus, the approach is consistent with the movement in the evaluation literature toward seeking alternatives to the limitations of the logic model, integrating complexity, and systems evaluation (Funnell & Rogers, 2011; Rogers, 2008; Hummelbrunner, 2011; Williams & Hummelbrunner, 2009).

The most powerful benefit of placing the PT in context is it assists the evaluator in managing the expectations of program staff and sponsors alike by showing those outcomes that can realistically be expected to change. When the targeted underlying conditions are placed in context, it is easy to see there are many factors contributing to the problem over which the program has no direct and immediate control to change. The context map also makes obvious the likelihood of seeing change reduce as you move along the continuum from immediate to long-term outcomes. While it is important to show the relationship between the activity and the intermediate and long-term outcomes, using the context map makes clear whether changes in longer term outcomes should be expected or promised.

Acknowledgments

The authors would like to acknowledge Dr. Carlos Rodriguez and Mrs. Gabriele Bartel for their insights and contributions to the manuscript. We are indebted to them.

References

Alkin, M. C. (2012). *Evaluation roots* ((2nd ed.)). Thousand Oaks, CA: Sage Publications.
American Lung Association. Freedom from smoking program. (2014, December 22). ⟨http://www.ffsonline.org/program-overview.html⟩.
Anfara, V. A., Brown, K. M., & Mangion, T. L. (2002). Qualitative analysis on stage: Making the research process more public. *Educational Researcher, 31*(7), 28–38.
Bamberger, M., Rugh, J., & Mabry, L. (2006). *Real world evaluation: Working under budget, time, data and political constraints*. Thousand Oaks, CA: Sage Publications.
Blamey, A., & Mackenzie, M. (2007). Theories of change and realistic evaluation: Peas in a pod or apples and oranges? *Evaluation, 13*(4), 439–455. http://dx.doi.org/10.1177/1356389007082129
Chen, H. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage Publications.
Coşkun, R., Akande, A., & Renger, R. (2012). Using root cause analysis for evaluating program improvement. *Evaluation Journal of Australasia, 12*(2), 4–14.
Donaldson, S. I. (2007). *Program theory-driven evaluation science: Strategies and applications*. Mahwah, NJ: Erlbaum.
Donaldson, S. I., & Lipsey, M. W. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. In I. Shaw, J. C. Greene, & M. M. Mark (Eds.), *The handbook of evaluation: Policies, programs, and practices* (pp. 56–75). London, UK: Sage.
Fielding, N., & Fielding, J. (1986). *Linking data*. Newbury Park, CA: Sage.
Foltysova, J. (2013). *Validation of reconstructed program theory*. Tucson, Arizona: University of Arizona.
Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models*. San Francisco, CA: Jossey-Bass.
Hummelbrunner, R. (2010, November). Changing institutional approaches to using logic models. Paper presented at the 24th annual conference of the American Evaluation Association, San Antonio, TX.
Hummelbrunner, R. (2011). Systems thinking and evaluation. *Evaluation, 17*, 395–403.
Huntington, C., & Renger, R. (2003). Response to the OMB performance and management assessment of the health professions program. Testimony submitted to the Interdisciplinary Committee on Community Based Linkages reporting to HRSA.
Jefferson University Hospitals. JeffQuit—smoking cessation program. (2014, December 23). ⟨http://hospitals.jefferson.edu/departments-and-services/jeffquit/⟩.
Kansas Department of Health and Environment. Kansas tobacco use prevention program. (2014, December 22). ⟨http://www.kdheks.gov/tobacco/index.html⟩.
Legacy. Become an ex: A new way to think about quitting smoking. (2014, December 22). ⟨http://www.becomeanex.org/about-ex.php#⟩.
Leeuw, F. L. (2003). Reconstructing program theories: Methods available and problems to be solved. *American Journal of Evaluation, 24*(1), 5–20.
Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.
Mark, M. M., Henry, G. T., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs*. San Francisco: Jossey-Bass.
Morell, J. A. (2005). Why are there unintended consequences of program action, and what are the implications for doing evaluation? *American Journal of Evaluation, 26*(4), 444–463. http://dx.doi.org/10.1177/1098214005281354
Morell, J. A. (2010). *Evaluation in the face of uncertainty: Anticipating surprise and responding to the inevitable*. New York, NY: Guilford Press.
Morell, J. A. (2014, April). *On logic models and unintended consequences [Web log comment]*. Retrieved from http://aea365.org/blog/jonny-morell-on-logic-models-and-unintended-consequences/
Nguyen, T. D., Attkisson, C. C., & Bottino, M. J. (1983). The definition and identification of human service needs in a community. In R. A. Bell, M. Sundel, J. F. Aponte, S. A. Murrell, & E. Lin (Eds.), *Assessing health and human service needs: Concepts, methods, and applications*. New York: Human Science Press.
North Dakota Department of Health. Tobacco facts: ND quits. (2013). ⟨http://www.ndhealth.gov/tobacco/Facts/NDQuits.pdf⟩.
Padgett, D. K. (2008). *Qualitative methods in social work research* ((2nd ed.)). Thousand Oaks, CA: Sage.
Patton, M. Q. (2008). *Utilization-focused evaluation*. Beverly Hills: Sage Publications.
Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Thousand Oaks, CA: Sage.
Renger, R. (2011). Constructing and verifying program using source documentation. *The Canadian Journal of Program Evaluation, 25*(1), 51–67.
Renger, R. (2014). Contributing factors to the continued blurring of research and evaluation: Strategies for moving forward. *The Canadian Journal of Evaluation, 29*(1), 104–117.
Renger, R., Bartel, G., & Foltysova, J. (2013). The reciprocal relationship between implementation theory and program theory in assisting decision-making. *The Canadian Journal of Program Evaluation, 28*(1), 27–41.
Renger, R., & Bourdeau, B. (2004). Strategies for values inquiry: An exploratory case study. *American Journal of Evaluation, 25*(1), 39–49.
Renger, R., & Foltysova, J. (2012). Partnership in international research and education (PIRE) supplemental evaluation award: update report—the impact and outcome of the Amazon PIRE program. (October).
Renger, R., & Hurley, C. (2006). From theory to practice: Lessons learned in the application of the ATM approach to developing logic models. *Evaluation and Program Planning, 29*(2), 106–119. http://dx.doi.org/10.1016/j.evalprogplan.2006.01.004
Renger, R., Kidd, E., & Jansen, A. (2006). Evaluation options for prioritized objectives of the Arizona Comprehensive Cancer Control Plan (AZCCCP).
Renger, R., Passons, O., & Cimetta, A. (2003). Evaluating housing revitalization projects: Critical lessons for all evaluators. *American Journal of Evaluation, 24*(1), 51–64.
Renger, R., & Titcomb, A. (2002). A three-step approach to teaching logic models. *American Journal of Evaluation, 23*(4), 493–503.
Rogers, P. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation, 14*(1), 29–48.
Respiratory Health Association. Quit smoking. (2014, December 23). ⟨http://www.lungchicago.org/quit-smoking/⟩.
Rosas, S. R. (2005). Concept mapping as a technique for program theory development: An illustration using family support programs. *American Journal of Evaluation, 26*(3), 389–401.
Schalock, R., & Thornton, C. (1988). *Program evaluation: A field guide for administrators*. New York, NY: Springer Science+Business Media.
Stake, R. E. (2006). *Multiple case study analysis*. New York, NY: The Guilford Press.
Weiss, C. H. (1997). Theory-based evaluations: Past, present, and future. *New Directions for Evaluation, 76*, 35–46.
Williams, B., & Hummelbrunner, R. (2009). *Systems concepts in action: A practitioner's toolkit*. Stanford, CA: Stanford University Press.
W.K. Kellogg Foundation. W.K. Kellogg Foundation Evaluation Handbook. (2010). ⟨http://www.wkkf.org/resource directory/resource/2010/w-k-kellogg-foundation-evaluation-handbook⟩ Accessed 21.10.14.

Ralph Renger, Ph.D. is a professor in the Center for Rural Health at the University of North Dakota. During the last two decades he is responsible for evaluating numerous local, state, national and international programs. He specializes in advancing theory driven methods for program evaluation. He regularly contributes to national and international evaluation publications and conferences.

Dr. Foltysova is a Dr. PH graduate from University of Arizona, College of Public Health. She currently serves as an evaluation faculty in the Center for Rural at University of North Dakota. She also currently serves as a contributing faculty member in the Ph.D. Public Health program in the College of Health Sciences at Walden University. Her dissertation focused on advancing program theory reconstruction methodology.

Karin L. Becker is working on her Ph.D. in Communication and Public Discourse at the University of North Dakota. Her research interest is in health communication and her dissertation includes the intersection of provider-patient communication preferences and computer mediated communication to better articulate the impacts of chronic pain. She is an evaluator and researcher at the Center for Rural Health at the University of North Dakota School of Medicine and Health Sciences. She has over 13 years of teaching experience and currently teaches in the Communication Program at UND.

Eric Souvannasacd is a Ph.D. student in the department of Educational Foundations and Research at the University of North Dakota. His dissertation is focused on efficacy, community capacity building, and social justice work within higher education using auto-ethnographic methodology. He is also a professional staff member at the University of North Dakota Center for Rural Health. In this capacity he serves as an evaluator for both the Seven Generations Center of Excellence in Native Behavioral Health (SGCoE) and the Evaluation of Cardiac Systems of Care project.