

A Review of Program Theory and Theory-Based Evaluations

Dr. Glynn Sharpe
Nipissing University
100 College Drive
Box 5002, North Bay, Ontario
Canada

Abstract

The paper explores and defines what program theory entails and when and why it is appropriate to develop a program theory. Components of a program theory are explored with specific attention paid to theory-based evaluation. The process is a complicated and multi-perspectives are required to ensure that appropriate steps are undertaken which support program efficacy and its overriding success.

Key Words: Program Theory, Theory-Based Evaluation

1. Purpose of the Paper

Program theory and the evaluation of a program's theory have gained interest in the evaluation field. Multiple terms have been integrated into discussions of program theory development and evaluation including: program theory, theory-based, theory driven, and program theory evaluation (Rogers, 2000a). Regardless of the term used, it should be clearly stated at the outset of this paper that the content is provided with respect to developing a theory of the program works and that the evaluation of the program's theory is an evaluation of the *program* and not the theory. The purpose of the present paper is to present a brief review of the literature addressing the development and evaluation of a program's theory.

2. What is program theory?

It is commonly reported that the function of a program theory is to ascertain the theoretical sensibility of the program (Chen, 1990b; Lipsey, 2000; Reynolds, 1998; Rogers et al, 2000; Rogers, 2000a; Sedani & Sechrest, 1999; Stufflebeam, 2000; Weiss, 1997). A program theory consists of a set of statements that describe a particular program, explain why, how, and under what conditions the program effects occur, predict the outcomes of the program, and specify the requirements necessary to bring about the desired program effects (Sedani & Sechrest, 1999).

3. When to develop a program theory

The primary stage to program development is the conceptual foundation. Once this has been established, the program theory can be used to develop outcome and intermediate goals. According to Prosvac and Carey (1997), this sequence of planning stages increases the chance of program success. Therefore, a program theory should be developed prior to the commencement of the program (Bickman, 1987; Prosvac & Carey, 1997, Rogers et al, 2000). It is highly advisable to develop a program theory prior to the start of any program. This is not often the case (Bickman, 1987; Reynolds, 1998; Rogers et al, 2000; Stufflebeam, 2000). However, even if the program is underway, it is important for a program theory to be developed. Therefore, program theories can be developed during the operation of the program (Rogers et al, 2000) or prior to evaluating a program (Bickman, 1987). The development of a program theory is necessary when hoping to determine why a program is succeeding or failing and if and where program improvement should be focused.

4. Components of a program theory

Program theory modeling uses three components to describe the program: the program activities or inputs, the intended outcomes or outputs, and the mechanisms through which the intended outcomes are achieved (Reynolds, 1998; Rogers, 2000; Rogers et al, 2000; Sedani & Sechrest, 1999). A description of the critical inputs define the components of the program, describe how these components are delivered, define the strength or amount of treatment required to induce the outcome (Sedani & Sechrest, 1999), and outline the required aspects vital in producing the expected outcomes (Lipsey, 1993). The processes that the outcome is contingent upon (Lipsey, 1993) and that follow the inputs should be described.

These processes ensue during the course of participation in the program and ultimately contribute to achieving the desired outcome (Sedani & Sechrest, 1999). A detailed description of the process or mechanisms of the program theory include information about the important steps, links, and phases of the expected transformation process as well as some implementation issues. The output should specify the nature, expected timing, side effects, and pattern of change as well as interrelationships among outcomes (Chen, 1990; Lipsey, 1993, 2000; Sedani & Sechrest, 1999; Wholey, 1987). The output/outcomes can be broken into immediate, intermediate, and long-term impacts (Funnell, 2000). Implementation issues or resources necessary for carrying out the program's services (Bickman, 1987; Lipsey, 1993; Sedani & Sechrest, 1999) should also be detailed. For example, resources and implementation issues may include supplies, materials, and skills (Sedani & Sechrest, 1999).

5. Why develop a program theory

A program theory provides a basis for evaluating relatively uncontrolled programs. Specifying a program theory to planners, staff members, people responsible for obtaining funding, and evaluators will assist them to carry out their duties while explaining how funding is being utilized (Prosovac & Carey, 1997; Weiss, 1997). A program theory can also encourage program investors to be focused on specific outcomes, rather than wasting funding, resources, and measurement objectives on attempting too much (Prosovac & Carey, 1997; Rogers, 2000b;). The program theory clarifies the perspective of the program, on which an evaluation of the program's quality can be based (Bickman & Peterson, 1990). A program theory will supply a conceptual basis for refining and improving existing programs and also support inferences about new programs (Bickman, 1987; Lipsey, 1993). Thus, a clear program theory that has been evaluated and deemed successful will afford policymakers the opportunity to implement similar constructs to other relevant programs (Bickman, 1987). This information is vital to the practice of program developers, existing programs, and evaluators knowing what does and does not work within a particular program. Doing so will allow similar services to thrive.

6. Theory-based Evaluations

Once a program theory has been established, the process of conducting a theory-based evaluation can commence. One purpose of conducting a theory-based evaluation is to test the model hypothesized to explain the program and the mechanisms utilized to reach the intended outcomes (Rogers, 2000a; Rogers et al, 2000; Weiss, 1997). A number of vital components to an evaluation must be investigated in order for the findings to be reliable, valid, meaningful, and interpretable. Prior to commencing evaluations, the intended purpose of the findings and the level of complexity required should be considered, as they will direct the purpose and intricacy of the evaluation. The proposed impact that the results will have on the program also requires attention. For example, complex models may be necessary for those who have decision-making power but little background information on the program (Rogers, 2000a). Once the level of detail required in the evaluation has been determined, the evaluation can be conducted. The evaluator should consider variability in all aspects of the program including the clients, the causal mechanisms that include moderator and mediator variables, and observable outcomes and program effects (Lipsey, 2000).

The research design must be based on relevant constructs and variables, outcomes that occur prior to treatment and those attributed to treatment must be explored, and the overall theory must be interpretable and have practical implications (Lipsey, 1993). The program theory is vital in the theory-based evaluation; furthermore, the evaluation methodology requires careful consideration to determine whether the program, and which aspects of the program, are central in affecting change and for whom. Once the intermediate and outcome factors have been specified through the program theory, data collection can commence. Careful consideration the data collection instruments must be conducted and techniques for gathering information must be developed and implemented. Evaluation of the intermediate stages of the program is particularly difficult. A researcher must determine how to represent the program or treatment (the independent variable) and the techniques for evaluation (Lipsey, 1993).

Multi-method approaches are commonly used for the task of documenting the implementation of the program (e.g. ethnography, surveys, ratings, observations, and interviews) (Funnell, 2000; Lipsey, 2000) and these often become the tools to measure the program process or intermediate variables. Formal measures can be constructed to evaluate the services being provided within a program (Orwin et al, 1998); however, in developing these scales, psychometric properties must be considered (Cook 2000; Lipsey, 2000) as relevant measures are vital for accurately representing the intermediate goals and program process. Moreover, when considering measures, it is important to take into account the tools ability to detect or be sensitive to change (Bickman, 1987; Lipsey 1993, 2000).

Without this, the results may be inaccurate or not reflective of the program's practices and outcomes. Outcome measures generally evaluate the social conditions that the program was hypothesized to change, requiring the evaluator to translate program goals into measurable outcomes (Lipsey, 2000). Therefore, prior to measuring the outcome, it is vital to investigate the structure and substance of the outcome domain that the program is expected to impact and to then develop or choose an appropriate measure for the effects of the intervention (Cook, 2000; Funnell, 2000; Lipsey, 1993, 2000;). Outcomes can be measured once the program, or treatment is complete; however, investigations on the effects at the post program stage and during follow-up periods have also been suggested (see Reynolds, 1998 for a review of confirmatory program evaluation). An evaluation allows an investigation into the hypothesized relationships that contributed to the program's functioning. The causal inferences proposed in the theory are strengthened if the empirical patterns of results are consistent with the hypothesized effects of the program (Lipsey, 1993; Reynolds, 1998).

Thus to measure implementation, intermediate variables, and outcomes, a number of tools may be utilized; however, consideration of their reliability, validity, and applicability must be greatly considered. Although these factors must be scrutinized when establishing data collection methods, other considerations pertinent to the evaluation must also be well thought-out. Data collection is not only required on the intermediate goals, the program implementation, the processes mediating affects, and the anticipated outcomes, but also on the characteristics of the target population, (Sedani & Sechrest, 1999). Program clients are not homogeneous; therefore, characteristics of clients may strongly influence specific program components and outcomes (Lipsey, 1993; Sedani & Sechrest, 1999). For example, differential exposure to the program, problem severity, motivation, and ability level are all examples of circumstances that can increase variability in the achieved outcomes (Lipsey, 2000; Sedani & Sechrest, 1999). Therefore, as is the case with clients, certain variables are too important to ignore and should be evaluated and incorporated into the analyses. The program components utilized by the client, the amount of treatment actually received, and the integrity of the services provided is important when considering theory-based evaluations (Lipsey, 1993, 2000; Sedani & Sechrest, 1999).

Given that outcomes may depend on the amount of treatment received, this knowledge, along with the previously mentioned characteristics of the clients, may impact the intermediate processes and outcomes. Lipsey (1993) noted that the integrity of the treatment should also be considered, thus a poorly implemented program or poorly trained staff may interact with the level of intensity or treatment received and impact the outcome. The predicted time of impact should have been specified in the program theory. A program theory should outline the timing or occasions for the measurement of variables (Lipsey, 1993; Sedani & Sechrest, 1999). Without this level of information, results and interpretations will be misleading. The program components utilized by the client, the amount of treatment actually received, and the integrity of the services provided are all important considerations in theory-based evaluations.

It is vital to design an appropriate evaluation of the program theory as faulty designs lead to inaccurate results such as no program effects (Bickman, 1987). The evaluator must explicitly outline the relationships among the variables included in the program theory and those relationships should be stated in testable hypotheses (Sedani & Sechrest 1999). However, the evaluator must also understand and recognize meaningful outcomes, as some are more important than others (Lipsey, 1993). Moreover, regardless of the findings (statistically significant or not), practical significance should be considered (Lipsey, 1993), requiring some understanding of the concept under investigation. To understand just what produced the results, the evaluator must be sure of the mechanisms of the program that were pertinent to the outcome (Lipsey, 1993), thus a good program theory is required. A number of vital considerations in conducting and interpreting the results of a theory-based evaluation have been outlined. Evaluators must consider the impact of clients to the program, utilization of services, exogenous factors, relationships and interrelationships among program components, the complexity of the program, as well as the design of measures, timing measurement, and the length of time required to conduct a responsible evaluation.

7. Conclusions

In closing it is important to note that a theory is developed and refined over time as additional evidence (through evaluation or other research) is discovered, similarly more refined theories will lead to additional evidence (Lipsey, 1993; Wholey, 1987). Therefore, the theory development and evaluation processes are not immediate, but require time and extensive energy. This paper outlined the process of developing a program theory and provided detailed information related to theory-based evaluations.

Overall, the paper provided the detail necessary to conclude that developing and evaluating a program's theory is a complicated and intricate process; however, it has also been shown that the benefits of this method of evaluation for the efficacy of program delivery far outweigh the time, money, and manpower required to accomplish a theory based evaluation.

References

- Bickman, L (1987). The functions of program theory. *New Directions for Evaluation*, 33, 5-18
- Bickman, L & Peterson. K. A. (1990). Using program theory to describe and measure program quality. *New Directions for Evaluation*, 47, 61-73.
- Chen, H. T. (1990a). *Theory driven evaluation*. Thousand Oaks, CA: Sage Publications
- Chen, H. T. (1990b). Issues in constructing program theory. Issues in constructing program theory. *New Directions for Evaluation*, 47, 7-18.
- Cook, T. D.(2000). The false choice between theory-based evaluation and experimentation. *New Directions for Evaluation*, 87, 27-34.
- Funnell, S, C. (2000). Developing and using a program theory matrix for program evaluation and performance monitoring. *New Directions for Evaluation*, 87, 91-101.
- Lipsey, M. L. (1993). Theory as method: Small theories of treatments. *New Directions for Evaluation*, 57, 5-38.
- Lipsey, M.W. (2000). Evaluation methods for social intervention. *Annual Review of Psychology*, 51, 345-375.
- Orwin, R G., Sonnefeld, L J. Cordray, D. S., Pion, G, M., & Perl. H. I. (1998). Constructing quantitative implementation scales from categorical services data: examples from a multisite evaluation. *Evaluation Review*, 22, 245-288.
- Prosavac, E. J., Carey, R. G. (1997). *Program Evaluation: Methods and Case Studies* 5th ed. (pp. 102-120). Upper Saddle River, NJ: Prentice Hall.
- Reynolds, A, J. (1998). Confirmatory program evaluation: A method for strengthening causal inference. *American Journal of Evaluation*, 19(2), 203-221.
- Rogers, P, J (2000a), Program theory: Not whether programs work but how they work. In D. L. Stufflebeam, G. F. Madaus, & Kellaghan, T, (Eds.) *Evaluation models viewpoints on educations and human services evaluation* 2nd ed. (209-233). Boston, MA: Kluwer Academic Publishers.
- Rogers, P, J. (2000b), Causal models in program theory evaluation. *New Directions for Evaluation*, 87, 47-55.
- Rogers, P, J., Petrosino, A., Huebner, T. A., & Hacsí, T. A. (2000). Program theory evaluation: Practice, promise, and problems. *New Directions for Evaluation*, 87, 5-13.
- Sidani, S., & Sechrest, L (1999). Putting program theory into operation. *American Journal of Evaluation*, 20(2), 227-238.
- Stufflebeam, D.L. (2000) Foundational models for 21st century program evaluation. In D.L. Stufflebeam, G.F. Madaus, & Kellaghan, T. (Eds.) *Evaluation models on educators and human services evaluation* 2nd ed. (33-83). Boston, MA: Kluwer Academic Publishers.
- Weiss, C. H. (1997). Theory-based evaluation: Past, present and future. *New Directions for Evaluation*, 76, 41-55.
- Wholey, J. S. (1987). Evaluability assessment: Developing program theory. *New Directions for Evaluation*, 33, 77-92.