ELSEVIER

Evaluation and Program Planning 30 (2007) 339–350

EVALUATION and PROGRAM PLANNING

www.elsevier.com/locate/evalprogplan

Semi-structured interview protocol for constructing logic models

P. Cristian Gugiu[a,*], Liliana Rodríguez-Campos[b]

[a]*Interdisciplinary Ph.D. in Evaluation, Western Michigan University, Kalamazoo, MI 49009-5237, USA*
[b]*Department of Educational Measurement and Research, University of South Florida, Tampa, FL 33620, USA*

Received 7 April 2007; received in revised form 6 June 2007; accepted 5 August 2007

Abstract

This paper details a semi-structured interview protocol that evaluators can use to develop a logic model of a program's services and outcomes. The protocol presents a series of questions, which evaluators can ask of specific program informants, that are designed to: (1) identify key informants basic background and contextual information, (2) generate logic model elements, (3) model program inputs, activities, outputs, and outcomes, (4) build a rational theory, (5) develop a program theory, (6) prioritize logic model elements, and (7) build a graphical or tabular logic model. The paper will also provide an example of how this approach was used to develop a logic model for a youth mentoring program. It is our hope and belief that with this interview protocol, novice evaluators will be able to generate comprehensive logic models like seasoned professional evaluators.

Keywords: Logic models; Semi-structured interviews; Interview protocol; SSIP

1. Introduction

Assessing planned services to evaluate their merit and worth (Mark, Henry, & Julnes, 1999; Scriven, 1991) is a challenging and time-consuming process. Without a well-defined model to guide the evaluation design, program managers run the risk of implementing an evaluation plan that does not focus on the most salient dimensions of the program (e.g., activities, outcomes, etc.) and thus, may develop and implement a poor evaluation strategy. As a result, they may run out of time, money or political support before they can contribute real value to the program. Logic models have been in use since at least the 1980s when they were introduced to help evaluators identify essential program activities, set appropriate outcomes and develop a plausible theory for explicating the association between program activities and anticipated outcomes (McLaughlin & Jordan, 1999). The challenge for novice evaluators is that while several approaches for constructing logic models appear in the literature (e.g., Renger & Titcomb, 2002; United Way of America, 1996; W. K. Kellogg Foundation (WKKF), 2004a, 2004b), these approaches do not provide a comprehensive list of questions for soliciting relevant information from key informants to construct a comprehensive logic model. Instead, they provide a small sample of questions intended to act as exemplars. The purpose of this article is to provide novice evaluators with an interview protocol they may use to develop comprehensive logic models like seasoned professional evaluators.

The benefits of utilizing interview protocols to collect data and formulate decisions have long been known in many professions. For example, interview protocols have been used to make psychiatric intervention decisions (First, Spitzer, Gibbon, & Williams, 1996; Kulic, 2005; Rogers, 2003), explore the factors that affect physicians' assessment of patients' alcohol consumption (Aira, Kauhanen, Larivaara, & Rautio, 2003) and screen law enforcement personnel (Varela, Scogin, & Vipperman, 1999). One reason for their popularity may be that by focusing the data collection activities, interview protocols have the potential to standardize data collection and reduce the tendency of premature closure of data collection (reaching a decision on the basis of incomplete data),

*Corresponding author. Tel.: +1 269 267 0471; fax: +1 269 387 5923.
E-mail address: crisgugiu@yahoo.com (P.C. Gugiu).

0149-7189/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.evalprogplan.2007.08.004

anchoring (focusing too heavily on specific information), primacy and recency effects (recalling the first and last items of information, respectively, with greater frequency), or confirmatory (searching for information, interpreting new and existing information, or avoiding contradictory information to confirm one's preconceptions) biases. Despite their popularity in other professions, however, no interview protocols for constructing logic models can be found in the literature.

The semi-structured interview protocol presented in this paper will delineate an extensive series of questions evaluators need to ask program informants before developing an evaluation plan. The protocol is designed to match certain questions to specific informants, reduce redundancy and maximize comprehensiveness. The questions presented in this protocol are designed to solicit information that identifies key informants; basic information about the program; contextual factors that may impact either the program or the evaluation; the inputs available for operating the program; the planned activities and products that will be provided to program clients; the size and scope of the activities or products delivered or produced by the program; the anticipated short-term, intermediate, and long-term outcomes; the program theory that coherently binds each of these areas into a causal theory of change; and the priority that should be ascribed to each logic model element. We have also attempted to integrate checks into our approach that will enable users to identify activities and outcomes that have a reasonable change of occurrence. Finally, although this paper may appeal to academicians, the target audience is practicing evaluators who wish to add another tool to their toolbox of evaluation skills.

2. An overview of logic models

At some point during an evaluator's career, they will find it necessary to develop a logic model, particularly if they are evaluating a federally funded program. Since the passage of the Government Performance Results Act in 1993, government agencies have been responsible for establishing performance goals, choosing indicators for measuring these goals, and reporting annually on the success of meeting these goals (Cozzens, 1997; Office of Postsecondary Education, 1998). As a result of these requirements, agencies pressure program evaluators to provide them with the information they need to report. Logic models provide a way of structuring an evaluation to address these requirements. Specifically, the purpose of a logic model is to provide stakeholders with a visual map or narrative description of how specific program components are related to the program's desired results (Renger & Titcomb, 2002). Logic models serve numerous functions, including assisting evaluators to focus the evaluation on the principal elements of the program, providing staff and other stakeholders with a common understanding of program services and goals, identifying a set of performance indicators that may be used to develop a monitoring system and summarizing performance for funders and decision makers (Fashola, 2001; McLaughlin & Jordan, 1999; Rogers, 2005). Some scholars (Chen, 1990, 2005a; Coffman, 1999) have also argued that a logic model should incorporate the underlying program theory—rationale for why a set of actions might resolve a problem or produce a desired outcome—upon which the program is based to investigate its validity. Therefore, the ability to develop logic models that are in fact "logical" is a skill that evaluators need.

2.1. W. K. Kellogg foundation logic models

According to the W. K. Kellogg Foundation (WKKF, 2004a, 2004b), there are at least three different types of logic models that are optimized for different purposes. The *Theory Approach Model* depicts the theory of change (i.e., program theory) that influenced the design and plan of the program and therefore, is well suited for identifying "how and why [the] program will work" (WKKF, 2004a, p. 9). This model seeks to address the following four questions in a manner that matches the interests of the funder so as to influence them to reach a favorable decision: What issues or problems does the program seek to address? What are the specific needs of the target audience? What are the short- and long-term goals of the program? What barriers or supports may impact the success of the program?

Once a proposal is approved by the funders, program implementation may then begin. This task will require the breaking down of each activity into its constituent steps, the development of a timeline for each step, the implementation of a monitoring system to track progress, and the generation of solutions for obstacles encountered along the way. According to the WKKF, an *activities approach model* should be used for these situations because it links the planned activities together in a sequential order that maps the implementation process. As a result of the model's ability to monitor this process and detect implementation obstacles, it is a good tool to use when conducting a formative or process evaluation.

Finally, program directors may be asked to conduct an evaluation of their program as a condition for being funded. Typically, this will entail demonstrating that the program has produced positive changes in the target population. Occasionally, evaluators may be asked to demonstrate that the program has produced broader impacts on the organization, community or system. However, because the latter outcomes are distal in nature, funders will generally expect evaluators to investigate short-term (1–3 years) and intermediate (4–6 years) outcomes. Regardless of the scope of the evaluation, according to the WKKF an *outcomes approach model* should be used because it is designed to connect resources to activities, and activities to desired results. Therefore, it not only can be used to monitor whether activities are being implemented, but it can also explore whether these activities produce the desired results.

2.2. United way of America logic models

Although the WKKF approach is a popular method for generating logic models, its origins can be traced to another equally popular method. With over 125,000 copies of its manual for measuring program outcomes in circulation, the United Way's (1996) approach is among the most popular methods for generating logic models. It delineates four principal components: *inputs* are resources (e.g., staff, facilities, equipment, training material, etc.) a program uses to achieve its goals; *activities* are the immediate product of inputs and refer to the actions and/or processes implemented by a program to affect change; *outputs* are the products of these activities and include the amount of services provided, materials distributed, people served, etc.; and *outcomes* are the benefits derived by program clients and are a result, directly or indirectly, of program activities and services. Typically, outcomes are subdivided into three categories: short-term, intermediate and long-term.[1] Unlike outputs, which measure "dosage", outcomes indicate a change between a pre- and post-activity condition, usually related to knowledge, skills, attitudes, values, behavior or status. Finally, a fifth component may be incorporated into a logic model whenever it is important to note constraints on the program. *Contextual factors* refer to constraints or conditions under which a program operates (e.g., laws, funding requirements, staffing issues, location of services, transportation problems, and hours of operation).

2.3. Antecedent target measurement approach

According to Renger and Titcomb (2002), the purpose of a logic model is to clarify and test the rationale underlying the relationship between program components and the needs the program seeks to address. To this end, they developed a three-step approach, called the antecedent target measurement (ATM) approach. The first step of the ATM approach is to identify all of the antecedent conditions of the problem by collecting information from experts on a problem's underlying rationale. Specifically, they recommended that during interviews with key stakeholders, evaluators ask "why" questions until interviewees are able to explicate the underlying rationale of the problem. This method resembles the method of *process tracing* or *backward reasoning* in which researchers trace each step in a process from the observed effect back to the causal agent while eliminating alternative hypotheses along the way (Bennett & George, 1997; Chen, 2005b; Mahoney, 2000). Renger and Titcomb also recommended that evaluators examine the literature to determine whether there is evidence to support the causal inferences made by key stakeholders and to identify other causal factors that may have been omitted. Step two is to identify the antecedent conditions targeted by program activities. These conditions may be identified from examination of detailed program descriptions or protocols, observation of program services and interviews with program staff and clients. To distinguish program components from the list of antecedent conditions generated in step one, Renger and Titcomb suggested shading the program components and limiting the list of program components only to the ones the evaluator is directly responsible. Similarly, they advised against including resources in the logic model because they believed that "it is the responsibility of those intending to fund and implement a program to determine whether adequate resources exist to implement the program" (p. 500).

The final step of the ATM approach is to determine whether the outcomes are reasonable to include in the logic model in light of the evaluation timeframe. Renger and Titcomb observed that there is little value in committing resources toward identifying, implementing and monitoring outcomes that are not expected to change within the course of the evaluation. Moreover, they argued that if the causal links between short- and long-term outcomes are strong, one might logically assume that if short-term outcomes are observed, long-term outcomes will follow. Therefore, outcomes should only be included in the logic model if they are the goal of a program activity and there is a reasonable chance they can be observed to change within the course of the evaluation.

2.4. Open systems model

Cohen and Kibel (1983) developed the open systems evaluation approach to shift focus away from the attempts of some evaluators to use logic models to validate causal explanations, such as the ATM approach and Chen's (1990, 2005a) program theory approach. The open systems Model is based on the fact that evaluations are rarely able to utilize experimental designs or to establish causal explanations (Julian, Jones, & Deyo, 1995). Instead, open system evaluations focus on establishing a collaborative partnership between the evaluator and program staff to achieve strategic objectives and to measure impacts.

Unlike the prior two models, Cohen and Kibel incorporated a hierarchical method based on the expected level of change. Changes that are temporary (e.g., willingness on the part of key stakeholders or decision makers to learn more about a program or temporary changes in knowledge, skills, attitudes, or behaviors) are defined as the first level. The second level of the hierarchy requires sustaining the prior changes and program buy-in from key stakeholders. The third level involves broader changes in individual or organizational practices that prevent the onset or reduce the severity of problems. The fourth and fifth levels entail observable or measurable changes in the behavior of target populations and changes in social indicators reflecting reductions in problems, respectively. However, Cohen and Kibel acknowledged that a single program is unlikely to produce upper-level (community-level) changes. Therefore, this model incorporates the notion that an individual program is one

[1] These outcomes are sometimes also referred to as proximal, medial and distal outcomes, respectively.

component of a larger comprehensive strategy designed to solve community problems.

Despite these noted differences, the method for constructing a logic model using the open systems approach is similar to the prior methods. The evaluator begins with a statement of the problem and proceeds to identify program activities and outcomes based on the problem and scope of the program. Naturally, activities are expected to logically lead to short- and long-term outcomes. However, although a theory of change clearly underlies this approach, the purpose of this method is to facilitate the development of an evaluation plan rather than to investigate cause and affect relationships.

3. Semi-structured interview protocol (SSIP) for constructing logic models

In light of the existing logic model approaches, one may legitimately question how much novelty another approach can contribute. Our motivation in developing the SSIP was to fill the gap with respect to the need for a prescribed set of questions that evaluators could use to construct logic models. While a semi-structured interview is a familiar method to most evaluators, a search for the terms "logic model" and "semi-structured interview" in 47 scholarly databases yielded only one article (i.e., Lal & Mercier, 2002), which did not include an interview protocol for constructing logic models. Therefore, we expanded our search to all articles that contained the phrase "logic model." Although this search netted a considerable number of articles, perusal of these articles did not reveal an extensive list of interview questions or an interview protocol. This gap in the literature is surprising considering the need for logic models and the likelihood that an interview is a common method for collecting the information necessary to construct a logic model.

It is important to note, we do not profess that our approach of interviewing key informants for the purpose of constructing a logic model is novel because it is not, nor is the framework, upon which the SSIP is based, original. In fact, the SSIP shares many features with the aforementioned logic models. For example, we have adopted the logical framework proposed by the United Way and WKKF (specifically, the outcomes approach model). We have also adopted the process tracing method that underlies the ATM approach. And finally, we have incorporated a hierarchical method similar to the one utilized by the open systems model. Our contribution to evaluators is to present a comprehensive list of questions derived from a review of the literature and our own practice. We have organized the SSIP into seven major sections: gathering basic program and contextual information; generating the logic model elements; organizing these elements into inputs, activities, outputs, and outcomes; eliminating poorly conceived elements; identifying a plausible theory of change; prioritizing outcomes, and constructing a graphical or tabular logic model. Although the primary purpose of the SSIP is to assist evaluators in collecting relevant information that can be used to formulate an evaluation plan, we have found that the process of utilizing the SSIP greatly contributes to program development—particularly in the case of emerging programs—because it allows program managers to redesign or add program components when it appears that existing components are not likely to produce the desired outcomes.

3.1. Identifying key informants to interview

Whether a newly designed or established program is being evaluated, evaluators must collect relevant information from multiple individuals. Interviewing key informants provides them with an opportunity to become involved with the evaluation, which increases the likelihood the findings will be utilized (Cousins & Earl, 1995; Fetterman, 2001; Greene, 1987, 2005; Guba & Lincoln, 1989; O'Sullivan, 2004; Patton, 1997). Therefore, an important ingredient in utilizing the SSIP is the identification of the key informants who should be interviewed. According to Scriven (2005), stakeholders are classified in one of three groups: downstream impactees (e.g., target population and their family), midstream impactees (e.g., program staff), and upstream impactees (e.g., community). The choice of which informants one should interview depends upon whether the informants is in a position to have information of value to the evaluation. The following four set of questions are designed to identify downstream, midstream, upstream, and key evaluation informants, respectively.

- *Please identify the prospective or actual targets of your program*—What population is the program designed to serve? Do you anticipate that the family and/or friends of this population will benefit from the services provided to the target group?
- *Please identify the staff that work or will work on the program*—Who are all the program staff, either paid or volunteer, that work on the project? Are there any unfilled positions? If so, what are these positions?
- *Please identify indirect program impactees*—What groups do you think will indirectly benefit from the services offered to the target population? Which political and advocacy groups stand to gain/lose the most from this evaluation? What decision makers, advisory committees, administrators, legislators, community organizations, or consumer groups may have a stake in this program?
- *Please identify the evaluation key stakeholders*—Who will be the primary consumers of the evaluation? Who will see, has a right to see, or should see the evaluation findings?

3.2. Identifying basic background and contextual information

The second step in logic model construction is to collect basic and contextual information about the program. The

basic background questions are purely descriptive in nature and may be gathered from existing documents to reduce the length of the interview. The contextual questions, on the other hand, attempt to identify the moderating or mediating factors that may affect the effectiveness of the program. The purpose of this step is not to generate logic model components but rather to gather information that can provide evaluators with greater understanding of the general purpose of the program and the potential obstacles that it faces. Below are a series of questions that evaluators can ask program managers and directors, in case this information was not available in existing program literature.

- *Describe the program to be evaluated*—What is the name of the program? When did it start? Who started it and why? Is it similar to other existing programs and how?
- *Please identify the purpose of the program*—What is the purpose or philosophy of the program? Do you agree with this purpose? What problem or set of problems is it designed to correct?
- *Describe the financial situation of the program*—Who finances the program and why? What is the total budget for the program? How long is the program guaranteed funding? Are financial resources distributed as a lump sum, periodically, on the basis of submissions of deliverables, or reimbursements based on submission of invoices with required documentation?
- *Describe the capacity of the program*—How many clients will the program be able to serve per week, month, quarter, or year? How long will clients receive services? What is the capacity for each program component/activity? What is the anticipated average caseload per service professional?

Every evaluation is conducted within a context that influences not only the scope of the evaluation but also the means with which it is operationalized. Therefore, it is important to know the external factors that may influence program results, either positively or negatively. While these contextual factors may not be under the control of the program, awareness of them will enable program managers to design program components that take these factors into consideration and allow evaluators to anticipate alternative hypotheses that may threaten the clarity of the evaluation findings. Unfortunately, this information generally cannot be gathered from program documents. The best sources for this information are the research literature and program staff, including managers and directors. Below are a series of questions that evaluators should ask program staff.

- P*lease identify any contextual factors that may affect the program or evaluation*—Are there unique events or circumstances that could affect the program in ways that might distort the evaluation findings? Under what conditions or circumstances do you think the program will work best? Worst?
- *Please identify any social factors that may affect the program or evaluation*—What organizational or community factors do you think will help or hinder the program from achieving its goals? Are social attitudes in the community supportive of the program? How does the program take into consideration different cultural perspectives of program participants?
- *Please identify any program settings that may facilitate or impede meeting the needs of clients*—Do you think program settings such as facilities, event scheduling, location, group size, transportation arrangements, child-care, etc. will have any effect on the program? If so, what effect do you think they will have?
- *Please identify any pertinent legislature that bears importance on this program or evaluation*—Is this evaluation part of a broader government evaluation effort? If yes, what initiative is this evaluation part of?
- *Please identify any political factors and forces that could impact the evaluation*—What is the political climate surrounding the evaluation? What community groups or community leaders may contribute to the success or failure of the project? Explain how. What type of political pressures could the evaluation team encounter? From whom will these pressures come? What is their motivation and goal?
- *Please identify any controversy surrounding the program or evaluation*—Is there a controversy surrounding the program or evaluation? If yes, who are the proponents and opponents of the program and evaluation? What sparked the controversy? Have their views been considered by the program and evaluation? If no, why not?

3.3. Generating logic model elements

Logic model construction requires the adoption of a framework for categorizing the information collected. Thus, we adopted the framework proposed by the United Way and WKKF's outcomes approach model since no serious objections to this approach were found in the literature. However, these elements may be ordered in accordance with the preferences of the evaluator. We determined that respondents feel more comfortable when the order is outcomes, activities, outputs, and inputs. However, logic model construction is an iterative process that requires numerous passes to capture all relevant program components and desired outcomes. Therefore, one should expect to examine the results generated by these questions several times since responses to later questions will often require modification of previous responses (and no rearrangement of the question order will avoid this).

Program outcomes refer to changes (such as pre-/post-test differences, inter- or intra-group differences) that occur after program services are administered to the target population and may represent positive and negative changes or maintenance of a particular level or status that

would otherwise have deteriorated without program services. This category of results is further classified into one of three types of outcomes: short-term, intermediate and long-term. However, this division is a bit simplistic since the time periods are somewhat arbitrary. Therefore, we prefer to use a modified version of the hierarchical system devised by Cohen and Kibel (1983) in which outcomes are classified based on the level at which change is expected to occur. Short-term outcomes reflect temporary changes in knowledge, awareness, skills, attitudes, behaviors, performance, status, environment, or level of functioning, whereas intermediate outcomes reflect sustained changes in these domains. Long-term outcomes reflect organizational, community or policy level changes.[2] Consequently, evaluators should pay attention to identify goals at multiple levels.

Program activities are specific actions and processes used to produce outputs and outcomes. Traditionally, this logic model element primarily focused on addressing the question of what activities will be implemented. However, to gain a deeper understanding of the context within which each activity will occur, evaluators should also gather information on the intended target of the activity, who will implement the activity, and when and where the activity will occur. *Program outputs* refer to the direct results of program activities, such as services, products, techniques, tools, events and technology. Outputs are the preliminary results program managers hope will produce their anticipated outcomes. Typically, they are described in terms of the size and scope of the services or products delivered or produced by the program (WKKF, 2004a). Outputs address the questions of what services will be delivered or what products will be produced, to whom these services or products will be delivered or distributed, and in what dose or quantity they will be delivered or distributed in a specified period of time. Finally, despite the recommendation by Renger and Titcomb (2002) to exclude resources from logic models, resources are a critical ingredient in program development, operation and evaluation. Therefore, *program inputs* refer to all the resources invested and used by the program to achieve its outputs and outcomes.

3.4. Modeling program outcomes

An important step in composing a list of potential outcomes is to consider the multiple levels at which they may occur. A series of questions one may ask program managers and service providers are listed below. Our experience suggests that the first three sets of questions will generate the most responses from small service programs, while the remaining two sets are more appropriate for large programs designed to produce macro-level changes.

- *Individual- and familial-level*—What are the individual- or familial-level changes that may occur because of the program? What skills or knowledge will participants learn from the program? What changes in behavior or performance might one expect to see in program participants? What secondary benefits may family members derive?
- *Organizational-level*—What organizational changes may occur because of the program? What directions, career options, enhanced perceptions or improved skills may staff acquire? What service capacity may the organization develop or enhance?
- *Community-level*—What community changes may occur because of the program? What environmental changes may result from program activities? What social changes might one expect to observe because of the program? What economic outcomes could the program have on the local community?
- *System-level*—What specific system-level changes could the program have? What policies or legislative impact could this program have at the local or state level? What political impact could the program have if it is successful? Unsuccessful?
- *Statewide-, regional-, national-, and international-level*—What are the statewide, regional, national, or international changes that may occur because of the program?

3.5. Modeling program activities and outputs

Following the identification of short-term, intermediate, and long-term outcomes, the next step is to identify the activities the program intends to perform to achieve these outcomes. Similar to the structure utilized for identifying outcomes, evaluators should look for activities ranging from the micro- to the macro-level. Moreover, our experience suggests that evaluators should use the aforementioned question-to-program matching process to limit the length of the interview. Following the identification of program activities, evaluators should determine the quantity in which these activities will be delivered:

- *Individual- and familial-level*—What new or existing activities will the program provide to program clients or their families? When and where will these activities take place? Who will conduct these activities? Will clients be referred for any services? What client needs are these activities designed to meet?
- *Organization-level*—What new or existing activities will the program provide to staff? When and where will these activities occur? Who will conduct these activities? What staff needs are these activities designed to meet?
- *Community-level*—What new or existing activities will the program provide to the community? When and

[2]While the terms short, intermediate, and long are not completely descriptive, they reflect the natural progression one is likely to observe. That is, assuming program activities produce any changes, temporary changes will follow closely after program activities, sustained changes will require more time, and community-level changes will follow last.

where will these activities take place? Who will conduct these activities? What community needs are these activities designed to meet?

- *System-level*—What new or existing activities will the program provide to policymakers? When and where will these activities take place? Who will conduct these activities? What policy needs are these activities designed to meet?
- *Statewide-, regional-, national-, and international-level*—What new or existing activities will the program provide to the broader statewide, regional, national, or international community? When and where will these activities take place? Who will conduct these activities? What needs are these activities designed to meet?

3.6. Modeling program inputs

The next step in logic model construction is determining the resources needed to generate and support program activities. While the list of potential resources that may be used to support an activity is vast, such a list may be organized in a number of broad categories. Because responses to the questions below require intimate knowledge of program activities and the resources necessary to implement them, these questions are generally asked of program directors. Respondents should be asked to include volunteer and in-kind services that are donated to the program and unpaid overtime worked by program staff. While these inputs do not consume program resources, they must be known so that one can estimate the true cost of the program and thus, improve the accuracy of cost analyses. Finally, evaluators should also ask program directors if they have enough resources to implement and operate the program:

- *Resources*—What resources (facilities, equipment, materials, personnel, money, and other resources) are available to generate or support each of the aforementioned activities? May the evaluation team obtain a copy of the program's budget plan?
- *Resource gap*—Is there a gap between the resources necessary to operate the program and the available resources? What is the size and nature of the gap? How will this gap be filled? If the gap cannot be filled, which program activities or components are in danger of being cut or curtailed?

3.7. Building a rational theory

The construction of a program model that is "logical" requires that all identified elements conform to a set of measurement standards. According to the WKKF (2004a, 2004b), outcomes should be SMART: specific, measurable, action-oriented, realistic and timely. Failure to screen for these standards may preordain an evaluation for failure. Outcomes that do not meet these four criterion should be eliminated from or receive lower focus in the evaluation. Furthermore, every so often, there will be occasions when respondents provide outcomes or outputs that are incredulous to the evaluator based on their knowledge and experience. In these cases, evaluators should *gently* challenge respondents to provide a rationale for their expectations.

For example, one of the authors once worked as an evaluator on a federal project that provided technical assistance to organizations that were awarded grants to either provide services that led to the reunification of children in foster care with their parents or reduced the time to adoption for children whose parents' legal parenting rights had been terminated. One of the problems often encountered was that the logic models that grantees submitted to the regional evaluation center contained unrealistic outcomes. For instance, several grantees indicated that their goal was to reduce the number of children in the foster care system. However, these same grantees indicated that their anticipated objective was to adopt/reunify less than ten children per year. While in the technical sense this reduction of children from foster care does, in fact, reduce the number of children in the foster care system, it has a negligible impact on either the state or national foster care system. Therefore, had these grantees not adopted more realistic goals, the evaluation would have been forced to conclude that they failed to attain their principal outcome.

This anecdote hopefully not only cautions evaluators against accepting unrealistically high goals, but also against accepting meaningless goals. That is, had the organizations set a goal of reunifying or adopting only three children a year—a very low goal considering the size of the award each grantee received—the author would have needed to persuade them to raise their goal else run the risk that the final evaluation report submitted to the federal project officer include the statement "while the grantee achieved the goal they set out, the outcome does not appear to be worth the cost, particularly when compared with other organizations that attained more impressive results."

Listed below are sets of questions that we have found to be effective in ferreting out poorly conceived outcomes. It is important to note that not all of these questions need to be asked of respondents because in many instances the responses to these questions will be self-evident. Additionally, the degree to which an outcome is realistic or measurable can be determined from a review of the literature or consultation with a content expert:

- *Outcome is realistic*—What evidence is there to support that this outcome is attainable? Are you aware of any research that links program activities with this type of outcome?
- *Outcome is meaningful*—What difference will this outcome make in the lives of those who are impacted? How

will this improve the lives of clients, the community, etc.? Why is this outcome important? Will the outcome be worth the cost of the program?
- *Outcome is timely*—How long after having received program services is it reasonable to expect to observe the desired outcome?
- *Outcome is measurable*—Are there any existing instruments or methods for recording this outcome? Have they been used before to measure this outcome? If yes, what instrument or method was used? How successful was it in measuring the outcome?

3.8. Developing a program theory

Program theory provides meaning to the logic model by defining the connections among the previous four logic model elements. Program theory may be used for two purposes: (a) to determine the reasonableness of the rationale of how inputs support program activities that, in turn, meet client needs and produce desired outcomes, and (b) to form the basis for conducting a theory-driven evaluation (Chen, 1990, 2005b; Rossi, 1971; Weiss, 1972). This approach can be further disaggregated into two approaches. *Process theory* focuses on whether the program has taken the necessary steps to implement its planned services and activities, whereas *outcome theory* focuses on whether the theory of change, which forms the basis for why specific program activities are provided to the target population, is sensible:

- *Process theory*—Has a target population been identified? Are there adequate procedures for determining eligibility? Does the organization have adequate resources for supporting planned activities? Does the organization have the capacity to implement and operate the program? Do staffs have adequate educational credentials, training, work experience and supervision to perform the tasks that are expected of them? Is the current implementation plan adequate to meet future needs? Is there a monitoring system in place to assess the degree to which planned activities are implemented in accordance with expectations and needs?
- *Outcome theory*—Have the needs that underlie the problem of interest for the target population been identified? Do the planned activities meet the underlying needs of the target population? Are these activities offered in a high enough dosage to produce and sustain change in the desired outcomes? How will program activities produce the desired outcomes? What is the association between program activities and desired outcomes? Which program activities are most critical for attaining the desired results?

Investigating a program's process theory is critically important, particularly when one conducts a formative or process evaluation, because the insights gained from examining these questions can be used to help program managers improve the structure of their program and to more effectively plan for the future. Similarly, investigating the causal mechanisms that underlie the logic model elements is helpful but only to the degree to which it ascertains whether the logical links are plausible. Spending an excessive amount of resources to explicate the exact nature of program theory, on the other hand, is not judicious. This is not to say that there is no value in learning exactly how program elements are interconnected. It is just that this knowledge is superfluous to determining whether program services and activities produced a change in the desired outcomes, which can be determined using other means such as gain scores or process tracing. Therefore, we prefer to focus our attention on measuring anticipated changes and detecting unplanned side effects. However, as a check on the completeness and plausibility of our logic model, we review each element with our client to ensure that every long-term outcome has at least one intermediate outcome and that every intermediate outcome has at least one short-term outcome that precedes it, etc. To this end, we have found that reading the following instruction to the interviewee is a good way of engaging them in this process:
- *Result association*—For each logic model element, please indicate all the preceding and succeeding elements with which it is most likely associated and why. Please also note that in some instances an element may be associated with more than one element.

3.9. Prioritizing logic model elements

Next, we ask clients to prioritize each of the outcomes in order to determine how to best prioritize the evaluation resources. Specifically, we ask the program director or manager to distinguish which of the outcomes we should consider "critically important" to the evaluation of the program. Moreover, we instruct the interviewee only to designate an outcome as "critically important" if and only if they agree that failure on this outcome should result in failure of the entire program. Their response to this question not only informs an evaluator of where they should focus their attention, but it also allows program directors and managers to check whether they are focusing enough attention on activities they hope will produce the critically important outcomes. For example, in an evaluation of a mentoring program we forgot to define what was meant by "critically important." Consequently, the program manager indicated that more than half of the outcomes should be regarded as "critically important." After we realized our error, we asked the program manager if he was okay with the fact that we would have to fail the program (i.e., reach a negative summative conclusion or assign a failing grade to the program) if its performance on these outcomes fell below minimum acceptable standards.

Following this question, the manager revised his previous responses. In the end, he only regarded two outcomes as important enough to merit the failure of the entire program if the program's performance on these outcomes was inadequate.

- *Importance of result*—How important is each of the outcomes on a scale ranging from "critically important" to "not very important?" Please take great care in identifying an outcome as "critically important" to the overall evaluative conclusion because if your program fails or performs poorly on this outcome, the overall evaluative conclusion or grade given to the program will reflect this performance.

3.10. Building a graphical or tabular logic model

Once you have identified all the program elements, their relationship with each other, and determined which elements are critically important, the difficult part is done. The next step is to organize this information in a way that is clear to clients but still retains enough details to guide the evaluation. Traditionally, a graphical logic model is used to depict the logical flow and linkages between logic model elements, while a tabular or narrative approach is used to communicate greater detail. Most graphical models can be constructed by placing brief descriptions of each element into flowchart shapes that are arranged in a sequential order (inputs ⇒ activities ⇒ outputs ⇒ short-term outcomes ⇒ intermediate outcomes ⇒ long-term outcomes) and are connected by arrows that represent the causal links between elements. However, because logic models with more than 20 elements are too cluttered with arrows to be easily understood by clients, we have adapted a model proposed by Rodríguez-Campos (2005). Our model assigns a unique identification code at the top of each element and replaces arrows with the codes of the predicate elements, which are listed on the left side of the element. Furthermore, it utilizes borders to distinguish between critically, moderately, and minimally important activities and outcomes. Another significant departure from traditional logic models is the exclusion of program outputs. While information on outputs is collected and used to monitor program implementation, this information is omitted from the logic model due to its redundancy with program activities.

Fig. 1 presents part of the logic model constructed for the mentoring program. Although it illustrates a complex array of program activities and outcomes, it still lacks the necessary detail for developing a comprehensive evaluation plan. To this end, evaluators should utilize a tabular or

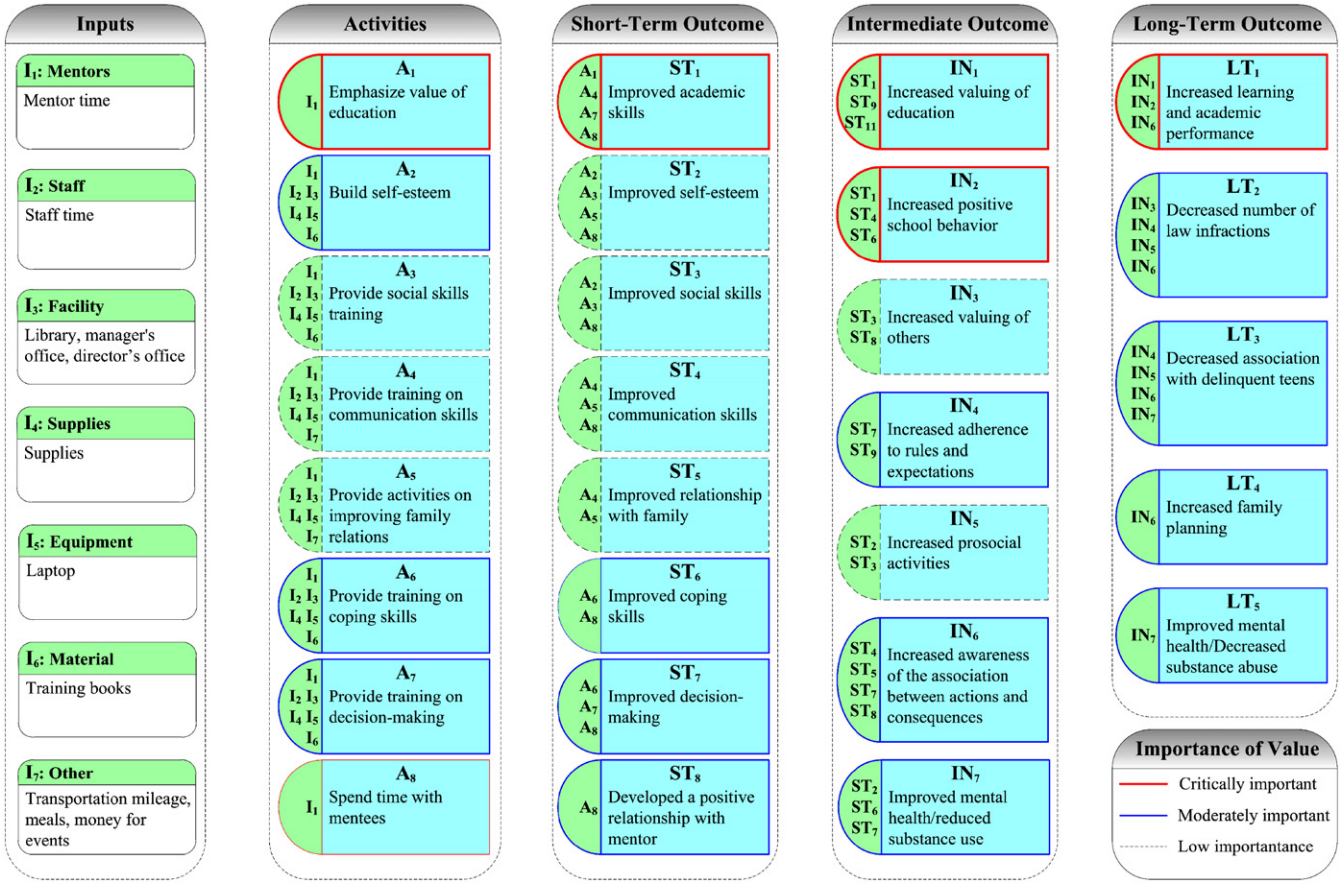

Fig. 1. Logic model for a mentoring program.

narrative logic model. A tabular logic model is similar to a graphical logic model in that program elements are organized into the same logic model structure. However, because this model is not constrained by space, evaluators may elaborate on each element. For example, evaluators can provide a short description of each element and its purpose, list indicators that will be used to measure performance on each element, specify sources from where data will be collected, etc. On the other hand, a narrative logic model provides narrative explanations of each element, the relationship between different elements, the underlying assumptions of the model and the indicators that are used to measure program implementation and outcome performance. Table 1 provides an example of a tabular logic model for the critically important logic model elements identified for the mentoring program.

Reflecting upon our approach we believe that it facilitated our ability to develop a comprehensive logic model for our client. First, the SSIP uncovered 34 potential outcomes (not all of which are presented in Fig. 1). Second, although identifying the program theory was the most time-consuming part of the SSIP to implement—because our client had just started the mentoring program and was not clear on their program theory—it generated several additional outcomes as they filled logical gaps in their program theory (e.g., a long-term outcome with no intermediate elements that linked the outcome to one or more activities). In addition to filling gaps in the program theory, the SSIP identified gaps in program activities that needed to be filled to improve the likelihood of attaining some of the outcomes. Third, the process of constructing the logic model uncovered information that led us to alert the director to several potential dangers (e.g., the lack of human capital to sustain the program in its third year). And fourth, the SSIP identified which program activities or outcomes were of critical importance. Consequently, we are able to focus more attention and resources on the elements that our client believes are important to the success of the program. As comparison, Table 2 presents a summary of the logic model that the program director constructed before we were contracted to conduct the evaluation. However, this comparison only highlights the difference between a logic model developed using the SSIP and one developed without the use of an existing approach or training in logic modeling.

Table 1
Logic model elements created by the SSIP

Logic model element	Data source
A1: Emphasize value of education	
A1.1: Mentors offer to check mentees' schoolwork	Weekly activity log
A1.2: Mentors advocate for child with teacher, principal, and parent	Weekly activity log
A1.3: Mentors teach mentees study skills	Weekly activity log
A1.4: Mentors teach mentees time management	Weekly activity log
A1.5: Mentors provide positive reinforcement for academic success	Weekly activity log
A1.6: Mentors explore mentee's academic problems and provide appropriate tutoring or referral	Weekly activity log
ST1: Improved academic skills	
ST1.1: Improved grades in school	School records
ST1.2: cc schoolwork with no prompting	Parent survey
ST1.3: Increased pride in academic ability	Mentee survey
ST1.4: Increased time spent doing school work	Parent survey
ST1.5: Increased pleasure derived from learning	Mentee survey
IN1: Increased valuing of education	
IN1.1: Increased GPA	School records
IN1.2: Completion of college applications	Mentee survey
IN1.3: Increased time spent formulating career plans	Mentee survey
LT1: Increased learning and academic performance	
LT1.1: Improved SAT/ACT Scores	Exam score card
LT1.2: Improved references for higher education	Mentee survey
LT1.3: Improved high school graduation rates for mentees	School records
LT1.4: Increased attendance of tertiary institutes, trade or technical schools	Mentee survey

4. Advantages and limitations of the SSIP

The SSIP approach has several advantages over the other logic model approaches. First, and foremost, it clearly lists the questions that evaluators should ask program staff and other key informants so that they can construct a comprehensive logic model. While other approaches certainly list a few questions (e.g., WKKF approach), these questions are only intended to act as

Table 2
Original logic model elements (not created by the SSIP)

Link 40 children with 40 mentors	
Outcome 1.1	All mentors have expressed interest in working with children in disadvantaged situations
Outcome 1.2	All mentors have completed screening and reference checks (i.e., abuse and criminal background)
Outcome 1.3	All mentors have received training and support in mentoring
Incorporate the elements of Positive Youth Development by providing youth with	
Outcome 2.1	Youth provided with safe and trusting relationships
Outcome 2.2	Youth provided with healthy messages about life and social behavior
Outcome 2.3	Youth receives guidance from positive adult role model
Outcome 2.4	Youth provided increased and enhanced participation in education for positive outcomes
Outcome 2.5	Youth participation in civic service and community activities
Outcome 2.6	Pro-social behavior will increase
Coordinate with partnering groups to develop plan for whole family	
Outcome 3.1	Support caregivers with training and help navigating services provided by the mentoring network
Outcome 3.2	Coordinate support services to siblings and families
Outcome 3.3	Connect the child with the imprisoned parent with permission from the other spouse or guardian

exemplars. Unfortunately, in our experience, the vast majority of individuals who serve as evaluators do not possess a graduate degree in evaluation and likely have not received extensive training on constructing logic models.[3] Consequently, they may benefit greatly from an interview protocol designed to assist them in constructing a logic model rather than the hit-and-miss of on-the-job training. Second, the SSIP incorporates questions designed to identify key informants; list potential contextual factors that may need to be monitored or considered; reveal the hierarchical level at which an activity, output, or outcome is anticipated to occur; eliminate poorly conceived outcomes; determine the importance of each outcome as judged by key program staff; and determine the hypothetical theory that underlies the logic model elements. While other logic model approaches certainly incorporate several of these elements, only the SSIP incorporates them all, to the best of our knowledge.

Despite the numerous benefits derived from logic models, they are not a panacea for program development or evaluation. Logic models are snapshots of current and planned program activities and desired outcomes. However, programs are rarely static. They change to accommodate new realities and interests. Consequently, logic model construction is an ongoing task rather than a one-time activity. Moreover, as Rogers (2005) pointed out, logic models that are excessively focused on intended processes and outcomes may lead evaluators to ignore the influence of other factors or to fail to consider unintended side effects. Therefore, while evaluators can use logic models to develop evaluation plans, these plans must retain enough flexibility to search for other potentially important outcomes.

Logic models are not always appropriate or necessary for conducting evaluations. Evaluation approaches that do not require knowledge of program goals—like Scriven's (1976) goal-free evaluation—do not necessitate the use of logic models. Moreover, logic model construction consumes considerable resources (Renger & Titcomb, 2002). For the mentoring program, 40% of the time spent developing the evaluation plan was devoted to the creation of a logic model (in total, approximately 35–40 h).[4] However, although evaluators may not need to organize the information outlined in this paper into a graphical, tabular or narrative logic model, this information must still be collected to conduct a program evaluation. Furthermore, although the cost of developing a logic model may be high, the process greatly enhances formative evaluations by uncovering weaknesses in the program implementation plan and focusing the evaluation on the activities and outcomes that really matter. For the mentoring program, the SSIP model led the program director and manager to focus more attention on providing actual mentoring activities (as compared with recreational activities). Moreover, a year after its construction, the logic model continues to act as a roadmap for all our monitoring and evaluation activities. Therefore, we believe that in the long run, logic models may actually be cost effective due to their ability to help evaluators focus their evaluative activities.

[3]Currently, there are less than a dozen doctoral programs in evaluation. Furthermore, our experience providing technical assistance to federal grantees suggests that the majority of individuals who served as evaluators on these projects lacked basic logic modeling skills.

[4]It is important to note that the mentoring program was in the formative stages of program development. Had the program been more mature, the interview process would have been shorter.

5. Final remarks

Logic model construction is an important first step in program evaluation. In our experience, very little guidance exists in the literature that outlines and organizes the questions one needs to ask of informants into an interview protocol. We believe that the SSIP approach outlined in this paper may be used by both novice and experienced evaluators to assist them in developing comprehensive logic models, which will improve the organization of their evaluations. Our seven-step approach begins with identifying key informants, background information, and contextual factors. Step 2 identifies key logic model elements. Step 3 organizes these elements into inputs, activities, outputs, and outcomes. Step 4 refines the emerging model by eliminating or reducing the attention paid to elements that are unrealistic, fall outside the scope of the evaluation, or cannot be measured. Step 5 explores and maps out the rationale that links the logic model elements. Step 6 identifies which elements are of critical importance to the program and evaluation. And finally, Step 7 utilizes the information from the previous steps to build both a graphical and a detailed logic model. While our approach is intended to be used in an interview format, it may be possible to utilize the SSIP as a survey distributed to key informants. This approach may be a fruitful area for future research for it may further reduce the cost of its implementation. We hope that evaluators will find our approach useful in the form presented here or at least be able to adapt it to meet their needs.

References

Aira, M., Kauhanen, J., Larivaara, P., & Rautio, P. (2003). Factors influencing inquiry about patients' alcohol consumption by primary health care physicians: Qualitative semi-structured interview study. *Family Practice*, *20*(3), 270–275.

Bennett, A., & George, A. L. (1997). *Process tracing in case study research*. Presented on October 17–19, 1997 at the MacArthur Foundation Workshop on Case Study Methods.

Chen, H. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage.

Chen, H. (2005a). Program theory. In Mathison (Ed.), *Encyclopedia of evaluation* (pp. 340–342). Thousand Oaks, CA: Sage.

Chen, H. (2005b). Theory-driven evaluation. In Mathison (Ed.), *Encyclopedia of evaluation* (pp. 415–419). Thousand Oaks, CA: Sage.

Coffman, J. (1999). Learning form logic models: An example of a family/school partnership program. *Harvard Family Research Project* [Online], available: 〈http://www.gse.harvard.edu/hfrp/pubs/onlinepubs/rrb/learning.html〉.

Cohen, A. Y., & Kibel, B. M. (1983). *The basics of open systems evaluation: A resource paper*. Available from The Pacific Institute for Research and Evaluation, 121 West Rosemary Street, Chapel Hill, NC 27516.

Cousins, J. B., & Earl, L. M. (1995). *Participatory evaluation in education: Studies of evaluation use and organizational learning*. London: Falmer.

Cozzens, S. E. (1997). The knowledge pool: Measurement challenges in evaluating fundamental research programs. *Education and Program Planning*, *20*(1), 77–89.

Fashola, O. (2001). Logic model basics. *Harvard Family Research Project*, *7*(2), 14–15.

Fetterman, D. M. (2001). *Foundation of empowerment evaluation*. Thousand Oaks, CA: Sage.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1996). *Structured clinical interview for DSM-IV axis I disorders, clinician version (SCID-CV)*. Washington, DC: American Psychiatric Press, Inc.

Greene, J. G. (1987). Stakeholder participation and utilization in program evaluation. *Evaluation Review*, *12*(2), 91–116.

Greene, J. G. (2005). *Stakeholder involvement. Encyclopedia of evaluation* (397pp.). Thousand Oaks, CA: Sage.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Thousand Oaks, CA: Sage.

Julian, D. A., Jones, A., & Deyo, D. (1995). Open systems evaluation and the logic model: Program planning and evaluation tools. *Evaluation and Program Planning*, *18*(4), 333–341.

Kulic, K. R. (2005). The crisis intervention semi-structured interview. *Brief Treatment and Crisis Intervention*, *5*(2), 143–157.

Lal, S., & Mercier, C. (2002). Thinking out of the box: An intersectoral model for vocational rehabilitation. *Psychiatric Rehabilitation Journal*, *26*(2), 145–153.

Mahoney, J. (2000). Strategies of causal inference in small-N analysis. *Sociological Methods & Research*, *28*(4), 387–424.

Mark, M. M., Henry, G. T., & Julnes, G. (1999). Toward an integrative framework for evaluation practice. *American Journal of Evaluation*, *20*(2), 177–198.

McLaughlin, J. A., & Jordan, G. B. (1999). Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning*, *22*, 65–72.

Office of Postsecondary Education. (1998). *Demonstrating results: An introduction to the Government Performance and Results Act*. Washington, DC.

O'Sullivan, R. G. (2004). *Practicing evaluation: A collaborative approach*. Thousand Oaks, CA: Sage.

Patton, M. Q. (1997). *Utilization-focused evaluation* (third ed.). Beverly Hills, CA: Sage.

Renger, R., & Titcomb, A. (2002). A three-step approach to teaching logic models. *American Journal of Evaluation*, *23*(4), 493–503.

Rodríguez-Campos, L. (2005). *Collaborative evaluations: A step-by-step model for the evaluator*. Tamarac, FL: Lumina Press.

Rogers, R. (2003). Standardizing DSM–IV diagnoses: The clinical applications of structured interviews. *Journal of Personality Assessment*, *81*(3), 220–225.

Rogers, P. J. (2005). Logic model. In Mathison (Ed.), *Encyclopedia of evaluation* (pp. 232–235). Thousand Oaks, CA: Sage.

Rossi, P. (1971). Boobytraps and pitfalls in the evaluation of social action programs. In F. G. Caro (Ed.), *Readings in evaluation research*. New York: Sage.

Scriven, M. (1976). Pros and cons about goal-free evaluation. In G. V. Glass (Ed.), *Evaluation studies review annual*, Vol. 1. Beverly Hills, CA: Sage.

Scriven, M. (1991). *Evaluation thesaurus* (fourth ed.). Newbury Park, CA: Sage.

Scriven, M. (2005). *Key evaluation checklist*. Retrieved August 19, 2005, from The Evaluation Center, Western Michigan University. Web site: 〈http://www.wmich.edu/evalctr/checklists/kec_april05.pdf〉.

United Way of America. (1996). *Measuring program outcomes: A practical approach*. Alexandria, VA.

Varela, J. G., Scogin, F. R., & Vipperman, R. K. (1999). Development and preliminary validation of a semi-structured interview for the screening of law enforcement candidates. *Behavioral Sciences & the Law*, *17*(4), 467–481.

Weiss, C. (1972). *Evaluation research: Methods for assessing program effectiveness*. Englewood Cliffs, NJ: Prentice-Hall.

W. K. Kellogg Foundation. (2004a). *Logic model development guide*. Battle Creek, MI: W. K. Kellogg Foundation.

W. K. Kellogg Foundation. (2004b). *Evaluation handbook*. Battle Creek, MI: W. K. Kellogg Foundation.