Program theory evaluation: Logic analysis

Astrid Brousselle a,*, François Champagne b

a Department of Community Health Sciences, Charles LeMoyne Hospital Research Center, University of Sherbrooke, Canada
b Department of Health Administration, Public Health Research Institute of the University of Montreal, University of Montreal, Canada

ARTICLE INFO

Article history:
Received 2 September 2009
Received in revised form 8 March 2010
Accepted 24 April 2010

Keywords:
Logic analysis
Evaluation
Evaluation theory
Program theory evaluation

ABSTRACT

Program theory evaluation, which has grown in use over the past 10 years, assesses whether a program is designed in such a way that it can achieve its intended outcomes. This article describes a particular type of program theory evaluation—logic analysis—that allows us to test the plausibility of a program’s theory using scientific knowledge. Logic analysis is useful for improving the intervention or finding alternatives for achieving intended outcomes; it influences the choice of type of evaluation to conduct and strengthens the validity of subsequent evaluations.

The objective of this article is to present the methodological principles and the roots of this type of program theory evaluation. We illustrate two types of logic analysis with two actual evaluation cases. There are very few published examples of program theory evaluation. This article will provide evaluators with both theoretical and practical information to help them in conceptualizing their evaluations.

1. Introduction

“Programs are complicated phenomena, generally born out of experience and professional lore” (Weiss, 1998). When requesting an evaluation, stakeholders generally want to know if what they are doing works and how they might improve their intervention. Program theory evaluation can often provide that kind of information without mobilizing the research efforts of an effect analysis.

Weiss defines program theory as “the mechanisms that mediate between the delivery (and receipt) of the program and the emergence of the outcomes of interest” (Weiss, 1998). Logic modelling is generally presented in the evaluation literature as a way to open the black box to better understand finer causal mechanisms. Some sophisticated logic modelling approaches include evidence-based depictions of the causes of the targeted problem, to better appreciate the intervention’s potential impact (Renger & Hurley, 2006; Renger & Titcomb, 2002). While this is an important step in better understanding the action mechanisms of the intervention, it is still not enough. As evaluators, we should also question the validity of the intervention’s chain of action (validity of the means), and we should test the scientific plausibility of the program’s theory. In fact, it could be argued that program theory does not really reflect how the intervention produces the intended outcomes, but rather, stakeholders’ perceptions and beliefs, right or wrong, about the mechanisms that operate between the delivery of the intervention and the intended outcomes. The whole evaluation is then built on the consensus reached on stakeholders’ beliefs and perceptions. But what do we do as evaluators if these are incomplete or, worse, if they are wrong? Can we really build valid evaluations based on the prior analysis of a program’s theory that reflects what people think, but not what the intervention does (Chen, 1990a, 1990b)?

As we will argue here, testing the program’s theory before entering more deeply into the evaluation process can provide important insights into the validity of the intervention’s means of action while mobilizing stakeholders in a valuable exercise of reflection (Pawson, Greenhalgh, Harvey, & Walshe, 2005; Renger & Hurley, 2006; Page, Parker, & Renger, 2007). It can also help in choosing the type of subsequent evaluation that will best suit the context and the intervention’s characteristics. There is an important body of documented experience of program theory evaluation (Campbell, 1966; Chen & Rossi, 1987; Lipsey, 1989; Marquart, 1990; Smith, 1990; Weissman, Silver, & Dillman, 2002). Several approaches and methodologies have been explored (Bickman, 1987a; Conrad & Miller, 1987; Cook, 2000; Lipsey, 1993; Weiss, 2000). Yet, in most cases important dilemmas were raised and significant evaluation effort was expended (Weiss, 1997), which may help explain why program theory evaluation is not more systematically used. In this article, we describe our experience of a certain type of program theory evaluation called logic analysis (Champagne, Brousselle, Contandriopoulos, & Hartz, 2009a; Contandriopoulos, Champagne, Denis, & Avargues, 2000). Logic analysis is a program theory evaluation based on existing
knowledge that uses expert judgment and scientific literature reviews. We begin with a presentation of the conceptual, methodological, and historical aspects of logic analysis. Then we illustrate the discussion with two examples, logic analysis of the Research Collective on Primary Care and logic analysis of integrated services for persons with dual diagnoses of mental health and substance use disorders.

2. Logic analysis

2.1. Definition

Logic analysis is an evaluation that allows us to test the plausibility of a program's theory using available scientific knowledge—either scientific evidence or expert knowledge (Champagne et al., 2009a; Contandriopoulos et al., 2000; Rossi, Lipsey, & Freeman, 2004). First, a clarification: logic analysis is not logic modelling. The model we present here has its roots in the approach that has been developed and taught for the past 20 years at the University of Montreal. This new proposal is built upon our practical experience of this particular type of evaluation.

Logic analysis is useful for better understanding the intervention's strengths and weaknesses and for analyzing whether the intervention is designed in a way that can logically produce the desired results (Champagne et al., 2009a). Furthermore, it allows us to assess the strength of the causal link between the intervention and the intended effects. In other words, logic analysis is a precondition for conducting an effect analysis; even so, in the absence of such a chain, logic analysis can provide insights into the intervention's potential (although that potential in no way guarantees the effects will be achieved).

An intervention's finer causal mechanisms—that is, the links between resources, processes and results—need to be examined to ensure there is strong internal validity and potential applicability to other contexts. For program theory evaluation, understanding an intervention's causal processes can be likened, in terms of validity criteria, to case study analysis. Case study analysis is a robust research design with strong internal validity, when based on solid theoretical foundations (Yin, 2002). Further, results from case studies can be applied to other contexts, precisely because of the detailed understanding of causal relationships (Yin, 2002).

Logic analysis follows the same reasoning. It allows us to understand the subtleties of the causal path and validate the intervention's theoretical construction with the added advantage of applying sound scientific knowledge. If the intervention's theory is validated through the analysis, its model will have good internal validity and could be applied to other contexts identified during the analysis.

The evaluator will rarely be able to find an evaluation already done of an intervention similar to the one to be evaluated. Nevertheless, the causal mechanisms at work in the program might have been analyzed and documented in other evaluation projects or social science studies. For example, there are numerous interventions based on behavioral theory whose hypothesis, briefly stated, is that if clients are informed, they will change their behavior. This principle, while applied in very diverse fields, remains the same hypothesis of action across programs and has been widely analyzed in the scientific field. Using logic analysis, the evaluator could thus test the program's theory by capitalizing on existing knowledge not only in a specific field of application of this principle (prevention of HIV/AIDS, smoking prevention, etc.) but also by using cumulative knowledge on the principles of action (do people change their behavior after having been informed, and if so, under what conditions?), which have been extensively explored.

2.2. Types of logic analysis

There are two ways to conduct logic analysis (Davidson, 2000). The first is direct logic analysis, which allows us to determine if the intervention is designed in such a way as to achieve the intended outcomes. This analysis judges the intervention's design and adequacy, and also provides important insights into how it could be improved. It will help in identifying whether the conditions are present in the design of the intervention for obtaining the desired effects, and it will provide information on the components that, according to scientific knowledge, should be added to improve its impact. It is mostly a formative evaluation.

The second way is reverse logic analysis, in which the evaluator looks for the best ways to reach the desired outcomes. This can be helpful for identifying alternatives that will produce the intended effects, broadening the array of interventions that could be implemented. This kind of logic analysis is essentially summative as, while identifying alternatives to action, it can confirm or invalidate the intervention under study.

It is principally the perspective of the analysis, and hence the orientation of the evaluation, that differentiates the two types of logic analysis. The direction of the analysis (direct or reverse) will necessarily influence the nature of the information that will be gathered, each one offering different evaluation outputs. The choice of doing a direct or a reverse logic analysis will depend on the kind of information stakeholders would like to have and their interest in the evaluation project. Would they prefer to get information that will help them strengthen and improve their intervention, or would they prefer to be aware of other means of action for achieving their objectives? The evaluation outputs will be of different natures, as the perspectives of analysis are different. Furthermore, both types of logic analysis could be conducted sequentially. The choice of analysis perspective should be discussed in advance with the parties participating in the evaluation project.

2.3. Steps for conducting a logic analysis

There are three steps in conducting a logic analysis. They are the same for both direct and reverse logic analysis. The perspective of analysis does not affect the evaluation strategy.

2.3.1. Building the logic model

First, the evaluator uses logic modelling to represent the intervention's program theory and the basis upon which it is supposed to lead to the desired effects. This is done using diverse sources of data, including documents and interviews with key stakeholders. Given the extensive literature available on logic modelling, we will not describe here the various steps and methods for constructing the intervention model (for an example, see Renger & Titcomb, 2002 or Champagne, Brousselle, Hartz, & Contandriopoulos, 2009b). Once the model is built, the evaluator will select a few issues to be explored in detail in the logic analysis. Issues are selected according to stakeholders' interests (e.g. are we achieving the desired outcomes?) or implementation difficulties encountered (e.g. we would like to get there, but find it difficult to apply that measure). The idea is not to simplify the model, but rather to focus on particular themes of interest to the stakeholders.

2.3.2. Developing the conceptual framework

Next, the evaluator gathers information on the best ways of doing things, either by looking at the intervention's main components to see if the optimal conditions have been assembled to achieve the desired effects (direct logic analysis), or by uncovering alternative ways of achieving them (reverse logic analysis). This step should use existing scientific knowledge, i.e.,
expert advice or research evidence identified through a literature review (Rossi et al., 2004). There may be different fields of knowledge leading to various theories. The evaluator is not constrained to choose one theory, but can instead propose a knowledge synthesis, highlighting divergences between theories. Further, the objective here is not to develop a systematic synthesis, which would require a considerably more intensive effort of literature analysis, but rather to clarify stakeholders’ representations using scientific knowledge. In drawing the conceptual model, not all perspectives are necessarily considered. Clearly, a trade-off might need to be made between the validity of the literature analysis and the time required to conduct an exhaustive analysis. Citing foundational and recent scientific work (without exhaustively consulting all the literature) or using evidence-based data such as systematic synthesis (or practice guidelines in the clinical field), if available, makes it possible to develop a sound representation of the intervention's theory.

2.3.3. Evaluating program theory

Finally, the evaluator will compare the intervention in action to the models that emerged from consulting experts or from the literature analysis. This should produce a new reading of the intervention that highlights its strengths and weaknesses, as well as the strength of the causal chain toward the effects. This step can be done by the evaluator, if he or she is familiar with the intervention, but also participatively, by collecting information through focus groups or interviews with key informants involved in carrying out the intervention. A participative approach will create consensus on what is required to improve the intervention, enhance the appropriation of results, and initiate the necessary changes (Donaldson, 2003).

2.4. Paradigmatic position

The logic analysis approach as described here fits within a postpositivist position, as it postulates that scientific evidence should orient the design of the intervention. In a constructivist perspective, scientific evidence should not be given as much importance as the various actors’ knowledge and representations. In logic analysis, the actors’ tacit knowledge should probably take the place of scientific evidence. However, we have not tested the applicability of such an approach using a constructivist perspective and adaptation would certainly be required.

2.5. The foundations of program theory evaluation

Logic analysis is a particular type of evaluation that belongs to program theory evaluation. In this section, we discuss its roots and what it has in common with other approaches to program theory evaluation and where it diverges from them.

2.5.1. Program theory in evaluation

Program theory has received a great deal of attention in program evaluation over the past 40 years. Among the foundational publications were special numbers of New Directions for Evaluation on the subject in 1987, and later in 1990 under the direction of Bickman, and in 2000 under the direction of Rogers and colleagues (Bickman, 1987b; 1990; Rogers, Hacsi, Petrosino, & Huebner, 2000b). This interest increased with the recognition that black-box evaluations were insufficient and that better knowledge of the theory underlying the program was necessary to produce generalizable findings (Bickman, 1987a). Further, the foundational work of Suchman (1967) and Weiss (1972) clearly influenced the field with the observation that failure to find program effects could, when not attributable to faulty evaluation design, be due either to the wrong theory or to inadequate implementation (Bickman, 1987b; Birckmayer & Weiss, 2000; Chen, 2004; Weiss, 2007). Implementation analysis addresses questions related to the implementation, and program theory evaluation addresses questions related to the adequacy of the program theory.

There is now a long list of terms associated with program theory, including theory-based, theory-driven, theory-oriented, theory-anchored, theory of change, intervention theory, outcomes hierarchies, program theory, intervening mechanism, theoretically relevant, logic modelling and program logic (Donaldson, 2007; Rogers, 2007), each term being defined differently according to the author who uses it. Different trends can be identified. Probably the three most important are logic modelling, evaluability assessment, and program theory evaluation. While each has different aims, they all developed from the initial recognition of the need to work on program theory.

Logic modelling is the representation of the program theory, which is, as Weiss defines it, “the set of beliefs that underlie action” (Weiss, 1998). Evaluability assessment was initially developed in the 1970s. It is a pre-evaluation activity that ensures programs are properly implemented and ready for summative evaluation (Trevisan, 2007). Evaluability assessment assesses the implementation’s compliance with the envisioned program (Fitzpatrick, 2002). Yet neither of these approaches addresses the plausibility of the program’s theory, which is a crucial difference from program theory evaluation, of which logic analysis is one type.

The idea of using a conceptual framework or program theory grounded in relevant substantive knowledge to guide evaluation efforts seems to have taken hold in the 1990s with another trend, called theory-driven or theory-based evaluation (Donaldson, 2007). A variety of approaches co-exist in this trend, all giving program theory an important place in the evaluation exercise. Among the most well-known and formalized approaches are Donaldson’s and Chen’s theory-driven evaluations (Chen, 1990a, 1990b, 2005; Donaldson, 2007) and Pawson’s and Tilley’s realistic evaluation (Pawson & Tilley, 1997, 2005). The sound objective of these approaches is not only to represent the program’s theory, but also to test it and understand the finer mechanisms that lead to the outcomes.

2.6. Program theory evaluation

Program theory evaluation helps “to specify not only the what of program outcomes but also the how and the why” (Weiss, 2000). The question of the intervention theory’s validity is all the more important, in that it is widely recognized that programs are generally poorly designed, being most often built according to the current stakeholders’ beliefs and not based on sound knowledge (Bickman, 1987b; Fitzpatrick, 2002; Shadish, 1987; Weiss, 1997). Programs are constructed over time within a process largely anchored in people’s patterns of functioning or designed according to their perceptions of what seems to work (Birckmayer & Weiss, 2000). Weiss stressed that social programs are complex and that they are “an amalgam of dreams and personalities, rooms and theories, paper clips and organizational structure, clients and activities, budgets and photocopies, and great intentions” (Weiss, 1998). She also found that many program theories were based only on practitioners’ assumptions and logical reasoning and hence were “simplistic, partial, or even downright wrong” (Rogers, 2007; Weiss, 1997, 2007).

The validity of a program’s theory can be assessed from two positions. First, we can do a judgment a posteriori. As stated by Suchman, Weiss and Scheirer, if an evaluation shows that the program was effectively implemented but it didn’t lead to the intended effects, the program’s theory must be invalid (Champagne, 2002; Scheirer, 1987; Suchman, 1967; Weiss, 1972). This is
a judgment made a posteriori. However, program theory evaluation should be conducted prior to other kinds of evaluation, in particular because of the influence it can have on the choice of type of evaluation.

The other position is to develop methodological devices to test the plausibility of the program's theory a priori. This approach is quite recent, and there are as yet few methodological examples. Historically, several authors have lamented the lack of a well-articulated approach to developing tests of program theory (Conrad & Miller, 1987; Mark, 1990). In 2000, Rogers and colleagues observed that, despite the apparent popularity of program theory evaluation, few examples existed in the formal evaluation literature (Rogers, Hacsi, Petrosino, & Huebner, 2000a).

Davidson and Hacsi added that, although evaluators increasingly recognize the added value of testing causal mechanisms, few actually do it (Davidson, 2000; Hacsi, 2000). Four years later, Rossi stressed the need for more work on how to test program theory (Rossi et al., 2004). Among the examples found in the literature, different methodologies were used to test program theory validity, the majority being experimental and quasi-experimental (Bickman, 1987a, 1987b; Conrad & Miller, 1987; Cook, 2000; Weiss, 2000) and others using observational and correlational studies (Lipsey, 1993). Among the methodological experiences, Chen and Lipsey used statistical analysis (correlational approach). Smith suggested path analysis, Campbell and Marquart used pattern matching to “compare the expectations generated by theory to empirical data to see how well they fit” (Weiss, 1998, p. 67) and Weitzman et al. used quasi experiment (Campbell, 1966; Chen & Rossi, 1987; Lipsey, 1989; Marquart, 1990; Smith, 1990; Weitzman et al., 2002). Weiss raised several limitations to these approaches. These methods can test the plausibility of only some of the hypotheses; however, in the case of complex programs, testing specific hypotheses is important rather than testing the program’s theory faithfully. Furthermore, if confronted with multiple program theories, the evaluator has to make a prior judgment on what theory to test (Weiss, 2000). Finally, complex programs require intensive evaluation efforts and huge amounts of data for tracking each small step of each causal chain (Weiss, 1997).

More recently, three models of theory-driven evaluation have developed explicit approaches and methodologies for testing program theory—those of Donaldson, Chen, and Pawson and Tilley. Donaldson identifies the active ingredients of the program and how they lead to desired outcomes (Fitzpatrick, 2002). He uses two major strategies to test the program: analyzing prior theory and research on the model, and analyzing the program while it is running (Fitzpatrick, 2002). Chen distinguishes between descriptive theories, which describe what causal processes are expected to happen, and prescriptive theories (also called change models), which describe what ought to be done, and he proposes theory-driven process and outcome evaluations (Chen, 2004, 1990a, 1990b; Weiss, 1997). “Theory-driven process evaluation systematically assesses how the major components of a program plan are being implemented in the field” (Chen, 2005, p. 173). In theory-driven outcome evaluation, the stakeholders’ theory will be tested by comparing the model to empirical observations of the program’s functioning. Realistic evaluation asks not “What works?” or “Does this program work?”, but rather, “What works for whom, in what circumstances and in what respects, and how?” (Pawson & Tilley, 2005). In realistic evaluation, the evaluator will seek context-mechanisms–outcome pattern configurations (Pawson & Tilley, 2005). The evaluator’s theories will be put to the test using the practitioners’ experience. The end product will be models indicating how, in the practitioners’ experience, programs activate mechanisms, among whom, and in what conditions, to produce the intended outcomes (Pawson & Tilley, 2005).

In which this approach is rooted have also been applied to research synthesis with a focus on better understanding and documenting, with evidence, how the intervention works (Pawson et al., 2005). This particular approach has been called realist review.

### 2.6.1. How logic analysis relates to other types of program theory evaluation

The three recent models of theory-driven evaluation we have mentioned—those of Donaldson, Chen, and Pawson and Tilley—can be divided into two categories. The first category documents the program’s theory and tests it according to empirical observations. This is the case in Chen’s approach and Pawson and Tilley’s realistic evaluation. The second category tests plausibility according to existing theories and research, which is what logic analysis does. This is the case in Donaldson’s approach and in realist review. Logic analysis is directly in line with Donaldson’s work on program theory evaluation, and specifically what he calls the plausibility check (Donaldson, 2007). According to Donaldson, the plausibility check should be the first step in the process of developing program impact theory. However, he passes over the subject rather summarily, such that further development would be needed to master it. The approach we present here should be interpreted as complementary to Donaldson’s description of his work under plausibility testing. Logic analysis also shares several points in common with realist review (Pawson et al., 2005). Realist review elicits the action theories that explain the principles underlying the context and mechanisms leading to attainment of the intended objectives. The principles and methodology described by Pawson and his colleagues could certainly be applied to direct logic analysis, which is consistent with their approach and objectives. The major difference between logic analysis and realist review is in the object of analysis of reverse logic analysis, which explores alternative ways of reaching the expected outcomes; this is not covered by realist review. The realist review approach, which has been well described and documented (Davis, 2005; Greenhalgh, Kristjansson, & Robinson, 2007; Mercille, 2008; Pawson et al., 2005), could be certainly be adapted to integrate the reverse logic analysis perspective, but more work would be needed along these lines.

### 3. Two examples of logic analysis

We used logic analysis twice, on different interventions. The evaluation of the Research Collective on Primary Care Services (Brousselle, Contandriopoulos, & Lemire, 2009; Contandriopoulos, Brousselle, & Kédote, 2008) illustrates direct logic analysis, while the logic analysis of integrated services for persons with dual diagnoses of mental health and substance abuse provides an example of reverse logic analysis (Brousselle, Lamothe, Mercier, & Perreault, 2007). Both interventions can be considered complex, as their implementation required significant social interactions. In addition, the case of the Research Collective on Primary Care Services illustrates the case of an innovative intervention.

#### 3.1. Evaluation of the research collective on primary care services in Quebec

#### 3.1.1. Program and issues

We were invited to evaluate the information use of an innovative experiment in knowledge building and dissemination: the Research Collective on the Organization of Primary Care Services in Quebec (Pineault et al., 2006, 2007). The Research Collective is a method of knowledge synthesis that promotes the transmission of knowledge and its use in policy development and in the management of organizations and systems. This method is based on a number of principles that distinguish it from systematic reviews. First, the collective integrates not only scientific
knowledge, but also the tacit knowledge of decision-makers. Second, in synthesizing knowledge, the collective bases itself on research studies in progress, rather than on published work. The objectives of this approach are to bypass the delays associated with getting research results published, to promote knowledge trans- mission in real time, and thereby to maximize the relevance of scientific knowledge in relation to the current context. The Collective was principally motivated by the fact that the organization of the news media services had then designated a priority area by the Ministry of Health, and many research projects on this theme had been funded. The synthesis was published as a report and then presented and discussed by the panel of decision- makers in a plenary session, before an assembly (n = 111) made up of researchers and decision-makers in nearly equal numbers (for detailed description see Pineault et al., 2006, 2007).

Stakeholders were interested in knowing if they had an impact on decision-making: was the information they produced used by decision-makers? Evaluating effects (information use) presented a number of challenges. How could we measure the different types of utilization (instrumental, conceptual, symbolic)? Could we be sure that the utilization as reconstructed was actually due to the Collective’s work rather than the utilization derived from the exercise? Faced with these challenges, we decided to proceed differently and conduct a logic analysis. We first documented the intervention using the grey literature and published articles on the Research Collective. We also had a meeting with the researchers of the Collective to clarify and build the logic of the intervention.

3.1.2. Development of the conceptual framework

The first current of literature used in designing our framework is derived from the conceptualization of the links between a information and decision-making in political science at large (Kingdon, 1995; Lindblom, 1959; Sabatier & Jenkins-Smith, 1999) and particularly in the subfield of lobbying (Ainsworth & Sened, 1993; Austen-Smith, 1993; Baumgartner & Leech, 1998; Berry, 1997; Carpenter, Estlering, & Lazer, 1998; Carpenter, Estlering, & Lazer, 2003; Carpenter, Estlering, & Lazer, 2004; Jordan, 1991; Milbrath, 1960, 1963). This literature provides convincing evidence that suggests that neither the content, the nature, nor the scientific quality of the information will have anything but a marginal impact on its utilization and thus that there is only a weak causal link between information transmission and its utilization in decision-making. The advantage of this first current is to highlight the notion that the core of the problem is competition for decision-makers’ attention, as well as competition between types of uncertainty and between sources of information.

A second current of literature—on evaluation—enriched the analysis framework with studies on the utilization of evaluation and, more broadly, of research (Beyer & Trice, 1982; Champagne, Lemieux-Charles, & McGuire, 2004; Patton, 1988; Weiss, 1977). These authors identify different categories of information use (instrumental, symbolic, and conceptual).

The third and last current of literature is the knowledge transfer and exchange (KTE) literature (Miron, Adair, McKenzie, Patton, & Waye Perry, 2007) as promoted by, among others, the Canadian Health Services Research Foundation (see for example CHSRF, 2003, 2005; Davis & Howden-Chapman, 1996; Denis & Lomas, 2003; Feldman, Gold, & Chu, 1997; Frenk, 1992; Lavis, Roberston, Woodside, McLeod, & Abelson, 2003; Lomas, 2005; Roos & Shapiro, 1999). A significant contribution of the KTE movement has been to suggest that the most relevant information, and therefore the most useful and usable, is a product of neither the politico-administra- tive nor scientific arenas, but rather a co-production made possible by the interaction of individuals in both (Denis & Lomas, 2003; Lavis et al., 2003). Likewise, the KTE school suggests that, although scientific information may generally be reliable, it can also be uncertain, dynamic, complex, debatable, and most often incom- plete (CHSRF, 2005).

The integration of these three fields approaching information utilization from different perspectives and using different concepts shows the complexity of the phenomenon (for a detailed presentation of the framework see Contandriopoulos et al., 2008). Because of this complexity, there could definitely be cases where scientific information of good quality is transmitted according to best practices without any subsequent utilization. Information utilization de- pending not only on the transmission process, but also, and primarily, on the contextual dynamics of the political arena, over which the researcher has no control—although adapting the strategy to the context is within the realm of the possible.

Three major dimensions seem to influence utilization: the context of the political arena, the traits of the information carrier and the characteristics of the information transmission process (Table 1). The only element over which the researcher has a certain amount of control is the transmission process. Convergences in the literatures on lobbying and knowledge transfer support the relevance and validity of these factors.

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimensions influencing information use.</td>
</tr>
</tbody>
</table>

1. Characteristics of the context
   - The system is open (deliberation (+) or closed (mediation of credible actors) (+)–) |
   - The topic of the information is on the agenda (+) |
   - There is political uncertainty surrounding the topic of the information (symbolic utilization (+) or programmatic uncertainty (conceptual and instrumental utilization (+) |

2. Traits of the information carrier
   - There is recognition/credibility of the carrier (+) |
   - There is proximity to the decision-maker (+) |

3. Characteristics of the process
   - There is significant personal involvement of the decision-maker (co-building of knowledge, knowledge brokering, etc.) (+) |
   - The information is submitted to a process of deliberation that includes different types of expertise: (+) according to the knowledge transfer movement; (–) according to the literature on lobbying |
   - Researchers and decision-makers are mutually useful (source of information/interest in the decision and decisional power) (+) |
   - The message is adapted to the needs of the targeted decisional environment (+), i.e., a formulated to convey the possible and probable consequences of decisions (+) according to the lobbying school; or b) formulated to provide advice on relevance, interests, needs, objectives, concerns, contextual information, and consequences (+) |
   - The main messages are formulated synthetically, concisely and attractively (+) |

4. Externalities related to the utilization of processes
   - The process may produce results of another type (creation of an exchange network, new alliances, etc.) |

(*) indicates a variable positively associated with utilization; (–) indicates the opposite; (+)–) indicates that no consensus exists in the scientific literature.
3.1.3. Evaluating the program theory

3.1.3.1. Data collection and analysis. We compared the various dimensions that, according to the established conceptual framework, would have had an impact on information use with the actual experience of the Research Collective. To do this, we analyzed the written documents available (reports, meeting notes, conference recordings, published scientific articles of the Research Collective) and conducted two group interviews, one with the Research Collective investigators and one with the stakeholders involved in the Collective. Interviews were recorded and transcribed. Data was analyzed by documenting each of the dimensions highlighted in the conceptual framework. We then were able to adapt the conceptual framework to the particular case of research findings utilization.

3.1.3.2. Results. The results of the analysis indicate that the Collective brings together many conditions that are favorable to information utilization: the window of opportunity presented by the disarray of the reform under way at that time and the resulting need for decisions at different levels of the health and social services network; the credibility of the information carriers; their proximity to certain decision-makers; the creation of a deliberative process involving decision-makers; the pursuit of consensus in crafting messages adapted to the decisional environment; and the formulation of these messages in formats that were summarized, concise, and attractive. Conversely, the utilization potential of the information was also reduced by various factors, of which most were outside the purview of the Collective. The central restructuration of the political uncertainty that characterized the decisional environment; the decision-makers' interests being channelled in accordance with the reform agenda; the difficulty of producing a message that was immediately transferable or adaptable to specific decisional needs, especially in an environment undergoing transformative changes; the divergent interests of researchers and decision-makers, or of decision-makers at different decisional levels; the pursuit of scientific neutrality despite the concern to ensure the messages would be useful for decisions; the limited representation in the process of certain categories of decision-makers; and the time constraints on the Collective's activities.

Logic analysis proved very helpful in evaluating information use in the case of the Research Collective. First, it confirmed the weakness of the causal link between the intervention (the knowledge transfer initiative) and the effects (the real uses of the information), which impeded the conduct of an effect analysis. As such, we decided not to pursue the evaluation. Second, because it is based on recognized scientific literature, we considered the conceptual framework to be robust enough to be applied to other knowledge transfer initiatives. We developed an innovative conceptual tool that could be used in other knowledge transfer contexts. We adapted the conceptual framework to the particular context of research knowledge transfer initiatives. Finally, this direct logic analysis allowed us to judge the potential impact of the Research Collective and showed that, given the contextual conditions, the Research Collective had greatly maximized the conditions for encouraging information use.

3.2. Logic analysis of integrated services for persons with dual diagnoses of mental health and substance abuse disorders

3.2.1. Program and issues

The prevalence of reported dual diagnoses is significant among patients in the mental health and substance abuse sectors, in addition to which many cases of dual disorders of mental health and substance abuse go undetected (Kédoté, Brousselle, & Champange, 2008; Mercier & Beaucage, 1997). Furthermore, because delivery systems generally do not have adequate resources to accommodate a clientele in need of multiple services, persons with dual diagnoses are often not well served. There is clearly a need to plan effective interventions adapted to the specificity of this clientele and its practice contexts.

Traditionally, the conflation between drug treatment and psychiatry have been profoundly isolated from each other, with separate histories and treatment philosophies (Brissin, 1988; SAMHSA, 2003). Guidelines for the treatment of dual disorders now exist, however, bridging the gap between both communities. It is recognized that dual disorders are better addressed using an integrated approach (Ziedonis et al., 2005) that is more than a simultaneous offer of services. “Integrated dual disorders treatments involve combining and blending the delivery of mental health and substance abuse interventions for persons with co-occurring disorders” (Torrey et al., 2002, p. 508). Dual diagnosis guidelines are largely inspired by the model of Mueser, Noordsy, Drake, and Fox (2003) that has proven its effectiveness.

Despite these guidelines, it remains difficult to integrate services for this clientele—not so much because of disagreements about clinical practices, but rather because of difficulties inherent in their organization. The integrated treatment model proposed by Mueser et al. (2003) seems to work if it is applied to the letter (Torrey et al., 2002). However, the faithful implementation rate remains low (Torrey et al., 2002), and the model is not necessarily suitable for all types of clientele, nor for several forms of health systems organization, in particular because of the required re-structuring of human and financial resources (SAMHSA, 2003).

In response to a need brought to our attention by professionals practicing in health institutions, we decided to evaluate, using reverse logic analysis, the model of service organization proposed by Mueser et al. (2003) and to see what alternatives might be identified beyond that model (Brousselle et al., 2007).

3.2.2. Development of the conceptual framework

Using reverse logic analysis, we tested the plausibility of the integrated treatment model's organizational hypothesis by collating and comparing the professional and the organizational literatures, and we proposed alternatives for the organization of care.

The professional literature deals with recommendations for dual diagnosis. The essential clinical and organizational components and the enabling factors put forward by Mueser et al. (2003) constitute what we called the professional model. We then analyzed the scientific literature on the dimensions of integration (Champagne et al., 2001; Contandriopoulos & Denis, 2001; Shortell, Gillies, Anderson, Erickson, & Mitchell, 1996; Sicotte, Léboux, Champagne, & Gamache, 2001; Touati et al., 2001), the organizational modalities of integration based on the work of Fleury (2002) and Page (2003), the centrality of organizations (Lamarche, Lamote, Bégin, Léger, & Vallières-Joly, 2002), the nature of organizations' links (Markham & Lomas, 1995), organizational autonomy, and the processes of network emergence (Conrad & Shortell, 1996; Glouberman & Mintzberg, 2001; Lamothé, 1999) (for details, see Brousselle et al., 2007).

By integrating the professional and organizational literatures, we were able to propose a conceptual model for dual diagnosis service integration that involves a reinterpretation of the current guidelines and offers new opportunities for service organization (Fig. 1).
3.2.3. Evaluating the program theory

3.2.3.1. Analysis. The conceptual model was built with input from a multidisciplinary team that included experts in evaluation, services integration, and dual diagnosis treatment. There was no need for specific data collection, but we presented our models to various treatment experts working with this type of clientele to discuss and validate our analysis.

3.2.3.2. Results. The construction of the conceptual model revealed important convergences between the organizational and professional literatures on factors predicting health services integration that highlighted the relevance of these factors in the specific context of dual diagnosis. Our analysis also enlarged the possibilities of integration models for the organization of care for dual diagnosis. The integrated treatment model proposed two principal forms of service integration: merger or cooptation. However, thanks to our analysis, these were no longer seen as the only solutions, and other forms could be envisioned. Professionals and managers were able to recognize the organizational dynamic they were involved in and to identify factors that played a critical role in service integration. They could consider how they might integrate services for this particular clientele in their specific context. Finally, our analysis repositioned the role of clinical guidelines: they were no longer an end in themselves, but rather became a starting point for health services integration.

Reverse logic analysis was a powerful tool for reinterpreting current guidelines in the field of dual diagnosis and opening new avenues for the organization of care for patients with dual diagnoses. In this case, after reverse logic analysis, we carried on with the evaluation of two completely different forms of innovative experiences of services organization for the same clientele by doing in-depth implementation analysis (process analysis) (Patton, 1997). The results have been submitted for publication (Brousselle, Lamothe, Sylvain, Foro, & Perreault, in press; Brousselle, Lamothe, Sylvain, Foro, & Perreault, accepted; Lamothe, Brousselle, Sylvain, Foro, & Perreault, submitted).

4. Discussion

Logic analysis proved to be an important evaluation to conduct before launching other evaluation activities. In fact, in both cases described above, it had important consequences for the various stakeholders, but also for the evaluation itself.

Let us consider the impacts of direct logic analysis on the evaluation of the Research Collective. First, our use of a deliberative strategy to discuss the conceptual framework and to assess the Collective’s impact potential with the researchers clearly contributed to their reflection on research knowledge transfer practices and on their role as researchers in the transformation of the healthcare system. Second, doing a logic analysis led us to reorient the type of evaluation we decided to do. Initially, the Collective wanted us to conduct an effect analysis. The logic analysis showed that the causal link between the intervention and the effects was too tenuous. Conducting an effect analysis in this case would have led us to flawed conclusions, either by measuring an effect and attributing it to the wrong element or by measuring no effect and concluding the intervention had no impact. A normative approach was more useful for determining whether the intervention presented the conditions that would maximize its potential impact (keeping in mind that even if the intervention includes all the conditions, there is no guarantee it will actually have an impact).

The third important contribution of logic analysis was to scientific knowledge, as it provided a better understanding of the factors that...
Influence information utilization not only in the general context of knowledge transfer, but also in the particular context of research. The reverse logic we conducted on the practice guidelines for dual diagnosis had similar consequences for stakeholders. A conceptual representation, for the choice of the type of evaluation to conduct, and for scientific knowledge. Usually, the simple representation of the intervention contributes to the enlightenment of the stakeholders. As Birckmayer and Weiss said, "In the end, whether or not the theory is right, it will have provided a framework for thinking about how the program is working" (Birckmayer & Weiss, 2000, p. 21). Yet, because the model is now based on sound knowledge, it opens the debate on solid foundations. In our case, clinicians and managers were able to recognize the patterns of their clinical practice, to envision the services integration process in their organization, and to identify reinforcement mechanisms.

In both cases, our logic analyses were conducted using existing scientific knowledge but we did not conduct systematic literature reviews, which would have required a far more extensive evaluation effort. With regard to the evaluation of dual diagnosis practice guidelines, logic analysis confirmed that it was appropriate to assess the interventions normatively according to the dimensions in the practice guidelines, but rather to develop a process analysis (Patton, 1997) that would integrate organizational innovations leading to alternative models of care that would bridge the psychiatry and rehabilitation sectors. Logic analysis then reinforced the appropriateness of the type of evaluation chosen. Finally, it is clear that, with logic analysis, we developed new scientific knowledge as we adapted knowledge on services integration to the domain of dual diagnosis of mental health and substance abuse.

As illustrated in both cases, logic analysis not only had an impact on stakeholders' appropriation of the intervention's conceptualization and triggered a reflection on their practices, but also had important consequences for the choice of the type of evaluation that appeared appropriate (or not) to do next and contributed significantly to the development of scientific knowledge.

The validity of logic analysis will depend on the evaluator's ability to identify the appropriate experts or literature domains. As we discussed earlier, there is a trade-off to be made between the time dedicated to the evaluation and the exhaustivity of the compiled information, unless the evaluator finds synthesized knowledge, such as in systematic literature reviews. Another challenge can be the integration of different theoretical perspectives. Often, the evaluation's perspective is influenced by the evaluator's original training. In logic analysis, the evaluator will be confronted with the existence of multiple theories in sciences that can produce a variety of perspectives that are each partially true but incomplete. It is not a requirement to use knowledge from various disciplines, but it may be too unrealistic to deliberately build a synthesis on a single field of knowledge. The evaluator might have to transcend disciplinary orientations and integrate the knowledge of diverse fields and reconstruct the “elephant” (as in the metaphor of the blind men and the elephant (Pawson & Tilley, 1997)) by combining all perspectives. Furthermore, when faced with a variety of what Pawson and his colleagues (Pawson et al., 2005) call explanation theories—that is, theoretical perspectives that may be contradictory in some of their aspects—the evaluator can build on these contradictions to identify tensions and potential unexpected effects. The objective of logic analysis is not to offer a final interpretation of the intervention but rather to bring some clarification and to stimulate reflection among the various stakeholders (Pawson et al., 2005).

Finally, logic analysis can be done in a relatively short time as compared to other types of evaluation, as it requires little data collection compared to other program theory evaluations that are based on empirical testing of causal relationships (Weiss, 1997). It can be carried out with few resources, and the output is quite enlightening for all participants. Logic analysis is useful for finding evidence on the conditions that maximize the interventions' effectiveness. It can have both a summative role, when assessing the validity of the intervention's means, and a role in formative evaluation, its clarifying function being helpful for modifying existing interventions or developing new ones. In the latter role, its impact is maximized if stakeholders have some latitude in redesigning interventions. Regarding direct logic analysis, the simpler and more specific the intervention's action mechanisms are, the easier it is to decide on the focus of the logic analysis. Further, if the intervention is based on well-known action mechanisms, it will be easier to find evidence to document their effectiveness and effects. Conversely, logic analysis will be more difficult to focus and to conduct if the intervention's scope of action is wide, i.e., not specific enough: it will require more extensive research. Further, if the means of action are based on very innovative action mechanisms, it might be harder to find evidence, which would render logic analysis less informative. Nevertheless, most of the time, even in very innovative interventions, the causal mechanisms at work are known and it is possible to find sound information on their validity and effectiveness.

These circumstances do not occur in reverse logic analysis, which is focused rather on finding the best ways to get the intended effects. Still, to facilitate the adoption of any new solution, the evaluator will have to use facilitation and prioritization techniques to help stakeholders determine the most appropriate means for their particular context of practice. Just as evidence-based medicine informs doctors and professionals on best practices according to scientific knowledge, logic analysis is a type of evaluation, rooted on realism approaches, that informs stakeholders on best practices for maximizing the effectiveness of interventions.

This article offers a first presentation of the methods, conceptual roots, and applications of logic analysis and it fills a need in a growing topic in the evaluation field (Weiss, 1997). As it becomes more extensively used, the usefulness and effectiveness of the approach will certainly also be more widely recognized.

5. Conclusion

In 1997, Weiss underlined the fact that theory-based evaluation “has generated considerable interest, but it appears to be having only marginal influence on evaluation practice” (Weiss, 1997). Thirteen years later, this is still the case. Her article raises several possible explanations for this phenomenon. Half are related to the difficulties of constructing the program theory due to the multiple theories of the agents or because of the hyperrealism of the modelling approach compared to the way programs and policies are conceived (Weiss, 1997). Even if these aspects still hold, they are now generally addressed in modelling techniques. The other half of the explanations is related to the difficulties of empirically testing the program theory's causal path. This is where this article makes its contribution. We do not propose empirically testing the causal path identified (which is possible but would involve mobilizing an extensive and intense research effort); rather, we propose using existing knowledge to test the program theory's validity. The advantage of theorizing interventions is that we can identify the more detailed causal paths that, most of the time, are well documented in the scientific literature or known by relevant specialists. Even innovative interventions are built on a combination of known and documented principles of action. These new methodological means of conducting program theory evaluation make the exercise more accessible to evaluation practitioners.
Regarding the issue of the generalizability of results, the ability to apply logic analysis results to other contexts will clearly depend, first, on the similarity of the action principles on which the interventions and the problem at hand are based. Scientific information used in the logic analysis and, finally, on the robustness of the information used. We have not gone into this discussion in depth because we cannot resolve here the debate on the value of scientific knowledge that has been launched by realistic and coheranian researchers (to cite only these two groups).

Clearly, this article presents one more step forward in developing methodologies for testing program theory. It does not solve all the difficulties that evaluators might encounter while conducting this particular type of evaluation, but we believe it offers several valuable insights and promising points for developing stronger intervention theories and evaluations.

Acknowledgements

The authors gratefully acknowledge the Fonds de Recherche en Santé du Québec and the Canadian Institutes of Health Research for their financial support.

References


