Theory-based evaluations: Framing the existence of a new theory in evaluation and the rise of the 5th generation

Astrid Brousselle
University of Victoria, Centre de recherche Charles-Le Moyne-Saguenay-Lac-Saint-Jean sur les innovations en santé, Canada

Jean-Marie Buregeya
Université de Sherbrooke, Centre de recherche Charles-Le Moyne-Saguenay-Lac-Saint-Jean sur les innovations en santé, Canada

Abstract
In this article we defend the idea that theory-based evaluations—contribution analysis, logic analysis, and realist evaluation—are complementary components of a new theory in evaluation. We also posit that we are currently observing the emergence of a fifth generation in evaluation: the explanation generation. Theory-based evaluations have featured prominently in the discourse of evaluators since the mid-1980s. They have developed mainly in response to the need for evaluation of complex interventions. In this article we analyze certain approaches that have matured in their design and application. We use the framework of Shadish et al. to analyze the ontological, epistemological, and methodological foundations of various theory-based approaches in evaluation to appraise their similarities and differences. We observe that all these approaches are grounded in critical realism. Similarities seen in their ontological, epistemological, and methodological positionings, as well as their complementarity in terms of the evaluative questions they address, suggest we may be observing the consolidation of a new theory in evaluation and the emergence of a fifth generation.

Keywords
critical realism, evaluation, program theory evaluation, theory, theory-based evaluation

Corresponding author:
Astrid Brousselle, School of Public Administration, University of Victoria, PO Box 1700 STN CSC, Victoria, BC V8W2Y2, Canada.
Email: astrid@uvic.ca

Introduction

In this article we defend the idea that theory-based evaluations—contribution analysis, logic analysis, and realist evaluation—are complementary components of a new theory in evaluation. We also posit that we are currently observing the emergence of a fifth generation in evaluation: the explanation generation.

By comparing various theory-based approaches on their positions in relation to knowledge construction, valuing, and use—dimensions used by Shadish et al. (1991) and Alkin (2004)—we see they share foundational principles and answer complementary, and even overlapping, evaluative questions in such a way that they appear to constitute a new evaluative theory. Theory-based evaluations, which have emerged in reaction to current "normal" evaluation practice, assert the need for a program theory when evaluating complex interventions. Their emergence marks a new *explanation* generation in evaluation that follows the four generations—measurement, description, judgment and pluralism—identified by Guba and Lincoln (1989).

Since the 1980s there has been an emergence and consolidation of theory-based evaluations (Blinded for review; Coryn et al., 2011). These are characterized by the development of a plausible program theory, which is a primary product of the evaluation and upon which are based results, recommendations, and conclusions. Evaluations of this type differ from logic model construction in their analytical focus: their objective is not to report on the expected links between resources, processes, and outcomes, but rather to provide a validated model that will enable a judgment to be made on the intervention being evaluated. Called theory-based, theory-driven, theory-anchored, or theory-oriented evaluations (Blinded for review; Coryn et al., 2011; Donaldson, 2007; Rogers, 2007), they are all aimed at reinforcing the explanatory power of evaluations (Weiss, 1997).

Program theory evaluation has historically attracted considerable interest in the field of evaluation. In the early days, this infatuation prompted Bickman (1987, 1990) and Rogers et al. (2000) to publish three issues of the journal *New Directions for Evaluation* in which the benefits, difficulties, and advances of these types of evaluations were discussed. Since then, various trends have been noted in the field of evaluation: an expansion in the number of evaluations identifying themselves as theory-based evaluations, an increased application of theoretically developed approaches giving prominence to program theory, and the growing body of literature presenting critical or synthetic analyses of theory-based evaluations (Coryn et al., 2011; Donaldson, 2007; Sridharan and Nakaima, 2012).

Since the 2000s, there has been a consolidation of approaches based on program theory. We can trace the methodological literature describing the approaches and their applications, as well as the literature that has strengthened the theoretical and methodological aspects by incorporating lessons learned from those applications. We will not here systematically analyze the practice of theory-driven evaluations, as has been done by Coryn et al. (2011); rather, we will analyze some of the approaches that have become more common over the past 20 years, to understand their foundations and compare them. Our objective is to attempt an in-depth reading of this movement to discern more clearly its core components and variations.

First, we present the methodological approach and its results by describing the selected evaluative approaches and where they stand in relation to the three foundations identified by Shadish et al. (1991): knowledge construction, valuing, and use. We then discuss the consequences of these results in terms of the consolidation of theories in the evaluation field.

Framework for comparing theory-based approaches

Chen (1990: 17) defined theory as "a frame of reference that helps humans to understand their world and to function in it." Program theory is defined as "a specification of what must be done to achieve the desired goals, what other important impacts may also be anticipated, and how these goals and impacts would be generated" (Chen, 1990: 43). In the literature, definitions, boundaries, and differences among program theory, intervention theory, and theory of change are not clear-cut (Funnell and Rogers, 2011; Weiss, 1997). As Weiss (1997) suggested "to keep the terminology simple", here we will use the term program theory to represent all causal pathways, including mechanisms and influences. Theory-based evaluations are based largely on program theory and are aimed at refining it using a variety of evaluative devices. A theory-based approach to evaluation represents "any evaluation strategy or approach that explicitly integrates and uses stakeholder, social science, some combination of, or other types of theories in conceptualizing, designing, conducting, interpreting, and applying an evaluation" (Coryn et al., 2011: 201). Program theory is the linchpin of the theory-based approach to evaluation; it describes how inputs, activities, and outputs centred on the program's process theory lead to immediate, intermediate, and final outcomes or impacts (Chen, 1990; Shadish et al., 1991).

Our objective is to analyze some of the approaches that have been conceived, applied, and adapted over recent years, and not to produce an exhaustive analysis of theory-based evaluations. Specifically, we will analyze three approaches—logic analysis, contribution analysis, and realist evaluation—by examining foundational articles, as well as articles that have proposed their refinement. Each approach will be analyzed in relation to the dimensions identified by Shadish et al. (1991) and taken up by Alkin (2004), which are knowledge construction, valuing, and use. Shadish et al. (1991) identified these as the foundational dimensions upon which all evaluation approaches should be assessed in terms of their positioning to understand in what ways they constitute a theory, as well as to appraise their similarities and differences.

The knowledge construction component addresses the issue of determining how evaluators construct reliable knowledge (Shadish et al., 1991). They approach this concept in relation to ontology, epistemology, and methodology. Ontology relates to positioning with regard to the state of reality of the subject being evaluated. It can be classified into three views: "1) the reality exists and is governed by immutable natural laws that are knowable; 2) the reality exists, but cannot be grasped objectively in all its complexity; and 3) the reality is a mental, social, or experimental construct, and is therefore multifaceted" (Champagne et al., 2011: 255, authors' translation). Epistemological positioning has to do with the relationship between the evaluator and the object of evaluation and standards of knowledge development. That relationship can be purely objective or subjective, or objective while incorporating contextual subjectivity, which could allow the use of evaluation results to be extrapolated to similar contexts (Champagne et al., 2011; Shadish et al., 1991). Methodological positioning refers to the technical devices and mechanisms used in the evaluation to develop knowledge (Champagne et al., 2011).

Valuing is central in evaluation. This component clarifies the evaluator's position regarding values (Shadish et al., 1991). Three objectives are possible: the development of a meta-theory – "the study of the nature of and justification for valuing" (Shadish et al., 1991: 48); a prescriptive theory; and/or a descriptive theory (Alkin, 2004; Champagne et al., 2011; Chen, 1990; Shadish et al., 1991). A meta-theory uses a reliable and structured method to substantiate the value's foundation and justification (Shadish et al., 1991). A descriptive theory

describes the values without necessarily assigning any supremacy (Chen, 1990; Shadish et al., 1991), whereas a prescriptive theory relates particularly to the supremacy of certain values when it indicates "what ought to be done or how to do something better" (Chen, 1990: 40). A prescriptive theory is characterized by a directed action, a formal conceptualization of the intervention, and an implementation strategy, as well as a range of options from which to select criteria related to outcomes (Chen, 1990).

The use component of evaluation theory refers to the utility of evaluative assessment in resolving social problems (Alkin, 2004; Shadish et al., 1991). This dimension is made up of three essential elements: intention, types of use, and ways of fostering use (Alkin 2004; Champagne et al., 2011; Chen, 1990; Shadish et al., 1991). The intention of the evaluation is formative when data are collected prospectively and the results are used to inform decision-making, the implementation process, quality assurance, and reporting. Evaluation has a summative role when information is collected retrospectively to determine whether or not the intervention should be repeated (Stufflebeam and Coryn, 2014). The intention is developmental when the objective of the evaluation is to foster the emergence of social innovation and the intervention's transformation in a dynamic environment (Patton, 2011). Although presented in the literature as exclusive categories, in practice we observe overlaps between these different intentions. They are more aptly considered to be abstract categories for analyzing approaches than exclusive categories into which evaluations fit. There are three types of use: conceptual, symbolic, and instrumental (Alkin, 2005; Champagne et al., 2011; Chen 1990; Patton, 1988; Shadish et al., 1991; Weiss, 1988). Conceptual use refers to changes in the conceptualization of a problem, of the intervention, or of potential solutions. Symbolic use is intended to demonstrate or legitimize certain positions that have been predetermined by the stakeholders. Instrumental use refers to real and programmatic changes related to the results of the evaluation, expressed in terms of decisions or their implementation. Ways of fostering use refer to all methods employed (Shadish et al., 1991), whether those are methodological processes to foster the transfer of results into other contexts, the commitment of stakeholders and their control over the evaluation process, or the means used to disseminate results.

Synopsis of selected approaches

Logic analysis is "a type of program theory evaluation that uses scientific knowledge to evaluate the validity of the intervention's theory and identify promising alternatives to achieve the desired effects" (Rey et al., 2012: 62). It is used to test the plausibility of the program theory (Brousselle and Champagne, 2011; Champagne et al., 2011). It sheds light on the program's strengths and weaknesses, elucidates the links between the program's design and the production of desired outcomes, and identifies contextual influences (Brousselle and Champagne, 2011; Contandriopoulos et al., 2015; Rey et al., 2012; Tremblay et al., 2013). Logic analysis examines: "(a) the important characteristics the interventions must have to achieve the effects and (b) the critical conditions required to facilitate the implementation and produce the effects" (Rey et al. 2012: 63). There are two types of logic analysis: direct and reverse (Rey et al., 2012). Direct logic analysis scrutinizes essential characteristics of the intervention and critical conditions leading to desired or other effects. It has important similarities with realist review, which documents the links between context, mechanisms, and outcomes using existing literature in the scientific and empirical fields (Pawson and Tilley, 2004; Pawson et al., 2005). Here we chose logic analysis because it

encompasses both direct and reverse logic analysis. Reverse logic analysis explores the best means to attain desired outcomes or other outcomes (Brousselle and Champagne, 2011; Rey et al., 2012). Logic analysis relies on available scientific knowledge, either evidence-based or expert knowledge (Brousselle and Champagne, 2011; Blinded for review; Contandriopoulos et al., 2008). Both direct and reverse logic analysis involve three steps: 1) building the logic model; 2) developing the conceptual framework; and 3) evaluating the program theory (Brousselle and Champagne, 2011).

Contribution analysis is an effect analysis approach (Mayne, 2001, 2008, 2011, 2012a). It examines, through credible causal claims, the contribution rather than the attribution a complex program is making to expected outcomes and impacts in complex settings (Delahais and Toulemonde, 2012; Mayne, 2011, 2012a, 2012b, 2015). Contribution analysis involves six key steps: 1) setting out the cause–effect issue to be addressed; 2) developing the postulated theory of change and risks to it, including rival explanations; 3) gathering evidence on the theory of change; 4) assembling and assessing the contribution claim and challenges to it; 5) seeking out additional evidence; and 6) revising and strengthening the contribution story (Mayne, 2012a: 272). It aims to "infer plausible association between the program and a set of relevant outcomes by means of systematic inquiry" (Lemire et al., 2012: 295). Dybdal et al. (2010) indicate that contribution analysis ascertains the program's contribution by establishing the postulated theory of change, identifying key threats to impacts pathways, establishing other contributing factors, and evaluating the principal rival explanations. It considers uncertainty in evaluating complex dynamic programs (Biggs et al., 2014). Five criteria with regard to the embedded theory of change must be met to infer causality for the program: "1) plausibility of the theory of change; 2) implementation as outlined in the theory of change; 3) evidentiary confirmation of key elements; 4) identification and examination of other influencing factors; and 5) the extent to which key alternative explanations have been disproved" (Mayne, 2011: 7). These steps can be supplemented by the use of the Relevant Explanation Finder (Biggs et al., 2014; Lemire et al., 2012), a tool that helps to clearly identify the factors influencing the chain of impacts, as well as alternative explanations.

Realist evaluation assesses complex programs by probing what works, for whom, and under what circumstances (Pawson and Tilley, 1997, 2004). Realist evaluation involves four core steps: 1) articulating the program theories to be tested, 2) collecting data to test the hypotheses; 3) testing the hypotheses; and 4) interpreting and refining them (Mehdipanah et al., 2015; Pawson and Tilley, 1997, 2004; Ranmuthugala et al., 2011; Salter and Kothari, 2014). It uncovers underlying implicit or explanatory theory leading to the program and its multiple components, and it identifies contextual factors that spearhead pathways of change to produce expected outcomes (Jagosh et al., 2015; Pawson, 2002; Pawson and Tilley, 1997, 2004; Ridde et al., 2012; Salter and Kothari, 2014). It is a logic of inquiry that illuminates the program theory underlying the inherent characteristics of program implementation (Hewitt et al., 2012; Pawson and Tilley, 1997, 2004) to investigate the generative mechanisms associated with the program (M), the contexts under which the pathways operate (C), and the ways in which outcomes occur (O) (Salter and Kothari, 2014). Context–mechanism–outcome (CMO) configurations, as outlined by Pawson and Tilley (1997, 2004), foster the examination of recurrent patterns in the midst of complex social reality through in-depth explanations of causal pathways. This helps the evaluator articulate the program theory to be investigated and test hypotheses to produce transferable advice based on that theory and to inform decisions as well as evidence-based policy-making processes (Hewitt et al., 2012).

Positioning of the approaches

Logic analysis, realist evaluation, and contribution analysis all have their foundations in the critical realism paradigm (Hewitt et al., 2012; Mayne, 2015; Pawson and Tilley, 1997; Pawson et al., 2004; Ranmuthugala et al., 2011; Rey et al., 2012; Salter and Kothari, 2014; Tremblay et al., 2013). Critical realism is founded on "ontological realism, epistemological relativism and judgmental rationality" (Groff, 2004: 10). The discussion presented here is largely inspired by the work of Bhaskar (2008) and other authors who built on his ideas.

Ontological realism considers the objects of knowledge to be: a) real structures and causal processes that "operate independently of our knowledge, our experience, and the conditions which allow us access to them" (Bhaskar, 2008; p. 15) and b) categorical distinctness between the transitive and intransitive domains of science. The transitive domain refers to lasting social structures as well as to causal pathways that might change over time, while the intransitive domain refers to our deeper construction of that reality. To apprehend the latter, ontological realism applies three layers: 1) the empirical, which is what can be observed or experienced; 2) the actual, which is what is known but cannot always be seen; and 3) the real, which is the hidden but necessary precondition for the actual and empirical (Walsh and Evans, 2014). To determine the worth or merit of an intervention, evaluators must make rational selections within existing knowledge to construct the program theory.

Epistemological relativism advances that "any piece of knowledge is (and must be) constructed by some individual person, for some specific purpose, and in some particular context; and its 'truth-value' can then only be determined relative to this purpose and context" (Quale, 2012: 106). Hence we can refer to it as trans-positional objectivity, which considers various possible positions towards a particular phenomenon (Sen, 1993). It is a continuum ranging from objectivism to subjectivism, as it incorporates subjective human values, perceptions, and judgments of reality. To understand the phenomenon under study, evaluators likewise need to take into account the multi-causal nature of social structures and grasp their complexity through the causal mechanisms leading to outcomes. They employ various types of data and knowledge, including quantitative and/or qualitative methods, which must fulfill the requirements for rigour within their field. The centerpiece in this endeavour is critical pluralism, inasmuch as it accounts for crucial characteristics and conditions of the phenomenon, which operates in an open environment. Critical pluralism advances that complex and wicked problems necessitate a plurality of methods and domains, which potentially encompass multiple perspectives, thereby taking into account both concrete and intangible varieties of causally efficacious systems and structures (Mingers, 2015).

Judgment rationality is a dimension not listed in Shadish et al.'s (1991) framework but identified as vitally important in Bhaskar's work. We added this dimension to ensure the production of reliable knowledge that relates "knowledge claims to real state of affairs, and in particular to underlying causal powers, grounded in the real essences of objects" (Groff, 2004: 87). To evaluate thus consists essentially in carrying out a value judgment of an intervention by implementing a sound methodology capable of supplying valid scientific and socially legitimate information on the intervention or any of its components (Champagne et al., 2011). This involves incorporating multiple perspectives from various actors, for whom the scope and sphere of activities may be different. Accordingly, the evaluator should maintain a certain distance with regard to knowledge construction and analyze the best case scenario meticulously to establish a value judgment on whether the intervention caused the observed changes or not.

All these approaches, while guided by natural laws that can be understood in a fractional way, apprehend reality as real even if it cannot be captured in great detail. An action is causal only when its effect is activated by a particular mechanism in a particular context (Pawson and Tilley, 1997). According to Pawson and Tilley, "this proposition—causal outcomes follow from mechanisms acting in contexts—is the axiomatic base upon which all realist explanation builds" (1997: 58). This principle underlies generative causation, upon which all the afore-mentioned approaches are based. Consequently, realism is the cornerstone of the realist evalu-ation (Jamal et al., 2015; Pawson and Tilley, 1997, 2004; Ridde et al., 2012; Salter and Kothari, 2014). Logic analysis, realist evaluation, and contribution analysis also have the power to explain unexpected outcomes by uncovering mechanisms hidden beneath the surface.

Logic analysis examines critical features of the main intervention or alternatives, as well as pillar conditions that might explain how the intervention leads to desired or other outcomes. Social interventions are socially mediated and conveyed by human agents, such that the trajec-tory from causal mechanisms to outcomes is nonlinear, unstable, and recursive. Logic analysis helps to uncover causal pathways that may be observable, or else discernible but not always perceptible, or even hidden. It uses multiple methods to gather knowledge from different domains and various types of data to show how the intervention causes social benefit and desired effects. It is able to shed light on unexpected results, such as how resistance to change impedes achievement of desired outcomes. Even if the foundational literature on logic analy-sis is not always clear, the arguments presented above, when carefully considered, support the notion that logic analysis is within the stream of critical realism. They also explain how an evaluator can grasp knowledge about causal pathways underlying the intervention.

As discussed above, contribution analysis assesses an intervention's contribution by taking into account other explanatory factors and rival explanations. Realist explanatory mechanisms operate in three ways: 1) by reflecting the intervention's embeddedness within the structured essence of social reality; 2) by providing statements that describe how both macro and micro mechanisms impact the intervention; and 3) by demonstrating how the intervention outputs are generated from stakeholders' activities and resources (Pawson and Tilley, 1997). "Contributory causes are called INUS causes: an Insufficient but Necessary part of a condition that is itself Unnecessary but Sufficient for the occurrence of the effect" (Mackie, 1965: 245). Contribution analysis is very useful for studying interventions nested in multilevel systems that are commonly characterized by uncertainties, turbulence, and dynamic mechanisms oper-ating in an open system (Mayne, 2015). The philosophy underlying contribution analysis is centred on critical realism, making it a powerful tool for assessing underlying assumptions and risks behind causal links and for identifying other key competitors, as well as unexpected results in the outcomes chain, inasmuch as it integrates the three previously mentioned layers of ontological realism to infer causal relationship between intervention and outcomes. Regularities in logic analysis and realist evaluation, as well as in contribution analysis, are the causal pathways that have been generated, which represent the chain of impacts and contex-tual factors that are likely to influence the production of effects (Easton, 2010; Mayne, 2012b).

All the approaches share the same **valuing component**. These approaches are prescriptive, in that they do not simply describe the causal mechanisms, but also make recommendations to improve the interventions' impacts. They also share some features of the **use component**. The intent of the evaluation is generally either summative or formative. The high level of explana-tion presented in the change models offers formative information to improve the intervention. Nevertheless, all three approaches could be used in a summative way. Their intended use types

are mainly either conceptual or instrumental. In terms of facilitating use, they engage multiple stakeholders iteratively to optimize and influence the use of evaluation results. They also incorporate different knowledge pertaining to various fields or sources and rely on theorization to enhance the transferability of results.

Whereas each approach addresses a different evaluative question, they all proceed in the same way: their research design is a stepwise approach that includes constructing program theories and then enriching and validating them based on knowledge derived from various sources of data and fields of knowledge. The fact that each approach addresses a different question explains why their main use intentions (formative, summative, or developmental) are different.

Table 1 shows how each approach is positioned in relation to the dimensions of knowledge construction (ontological, judgment rationality, epistemological, methodological positions), valuing, and use.

Framing the existence of a new theory in evaluation and the rise of the 5th generation in evaluation

The three approaches—logic analysis, contribution analysis, and realist evaluation—selected because they all have a strong focus on program theory, present one particular feature of interest: while they have similar bases for characterizing and differentiating evaluative theories, the evaluative questions they address are complementary. They are empirically tested approaches intended to respond to the main evaluative questions by: 1) assessing the plausibility of the intervention (logic analysis); 2) analyzing an intervention's effects (contribution analysis); and 3) analyzing the intervention's implementation (realist evaluation). However, the boundaries between them are not clear cut.

In fact, the three approaches have important overlaps that make them non-exclusive and complementary. For example, with respect to program theory, one can easily argue that direct logic analysis addresses mainly the resources–activities–outcomes mechanisms while contribution analysis puts the emphasis on documenting the outputs–outcomes–impacts chain, making these two approaches complementary components of the program theory. With a logic analysis further scrutinizing consequences as impacts, what would be the difference between a contribution analysis and a logic analysis? The answer is probably the stage at which the evaluation is done, with logic analysis staying at a more theoretical level, not documenting observed effects, whereas contribution analysis explains observed and documented effects of the intervention.

The same kind of argument could be made regarding differences between realist evaluation and contribution analysis. As both aim to document context–processes–effects mechanisms, aren't they similar approaches? It is quite obvious in reading about these approaches that they are different and offer complementary evaluation perspectives, but their differentiation is not as clear cut as we would hope, which is not surprising as, in the end, the objective is to build an evidence-based program theory.

A third example would be the differences and similarities between logic analysis and realist evaluation. Realist evaluation relies mainly on empirical data and observations; when the source of information shifts to written data, the type of evaluation done—realist review—is similar to direct logic analysis.

Table 1. Comparative positioning of logic analysis, contribution analysis and realistic evaluation.

APPROACH	KNOWLEDGE CONSTRUCTION						VALUING			
	Critical Realism			Methodology					USE	
	Ontology	Rationality	Epistemology	Question	Steps	Data sources		Intent	Intended types	Means to encourage use
Logic analysis	Ontological realism	Judgmental rationality	Epistemological relativism	Direct logic analysis: Is the intervention designed in such a way to achieve the desired effects (Rey et al., 2012: 63)?	1- Building the logic model	Documents+ interviews with key stakeholders	Prescriptive theory	Formative	Conceptual use	Involve different stakeholders
					2- Developing the conceptual framework	Expert knowledge + scientific literature		Developmental	Instrumental use	Integrative theorization approach
					3- Evaluating the program theory	Qualitative and/or quantitative data				
				Reverse logic analysis: What are the best ways to achieve desired effects (Rey et al., 2012: 64)?	1- Building the logic model	Documents + interviews with key stakeholders		Summative	Conceptual use	Involve different stakeholders
					2- Developing the conceptual framework	Expert knowledge + scientific literature			Instrumental use	Integrative theorization approach
					3- Evaluating the program theory	Qualitative and/or quantitative data				
Contribution analysis	Ontological realism	Judgmental rationality	Epistemological relativism	To what extent are observed results due to programme activities rather than other factors? (Mayne, 2008: 1)	1- Set out the cause-effect issue to be addressed	Documents + interviews with key stakeholders	Prescriptive theory	Formative	Conceptual use	Involve different stakeholders
					2- Develop the postulated theory of change and risks to it, including rival explanations	Expert knowledge + scientific literature				
					3- Gather the existing evidence on the theory of change	Qualitative and/or quantitative data		Summative		
					4- Assemble and assess the contribution claim, and challenges to it					
					5- Seek out additional evidence	Qualitative and/or quantitative data			Instrumental use	Integrative theorization approach
					6- Revise and strengthen the contribution story					

Table 1. (Continued)

APPROACH	KNOWLEDGE CONSTRUCTION						VALUING	USE		
	Critical Realism			Methodology				Intent	Intended types	Means to encourage use
	Ontology	Rationality	Epistemology	Question	Steps	Data sources				
Realist evaluation	Ontological realism	Judgmental rationality	Epistemological relativism	"What works for whom in what circumstances and in what respects, and how?" (Pawson and Tilley, 2005: 363)	1- Articulate programme theories to be tested 2- Collect data to test the hypotheses 3- Test the hypotheses 4- Interpretation and refinement	Documents + interviews with key stakeholders Expert knowledge + scientific literature Qualitative and/or quantitative data	Prescriptive theory	Formative Developmental Summative	Conceptual use Instrumental use	Involve different stakeholders Integrative theorization approach

As we have seen above, logic analysis, contribution analysis, and realist evaluation share several foundations related to ontology, judgment rationality, epistemology, valuing, and components of the use dimension. There are methodological differences in how they construct a valid program theory while responding to complementary but still different evaluation questions. Differences in their use are due to the fact that they address different types of questions. In our view, the coherence in the positionings of logic analysis, realist evaluation, and contribution analysis, combined with their complementarity, suggests strongly that these are the components of a new emerging theory in evaluation.[1] Clearly, further work on this topic would help to deepen the understanding of their links and complementarity in terms of evaluation questions. This work would lay the groundwork for describing this new theory, which represents a departure from the existing evaluative questions of plausibility, effect, and implementation analysis.

If we analyze the construction of theory-based evaluations from a historical perspective, the consolidation of a new evaluative theory might even mark the emergence of a fifth generation in evaluation: the explanation generation. In fact, Guba and Lincoln (1989) identified, in the history of the domain of evaluation, four generations of evaluation—measurement, description, judgment, and pluralism—the fourth being the only one marking a real paradigmatic shift. We would not go so far as to refer to paradigm change when describing the movement coalescing around the consolidation of theory-based evaluative approaches. Nevertheless, similarities in positioning around 'anomalies' is another element confirming the rise of a fifth generation. Here we observe the same phenomenon as for the first three generations: approaches sharing the same paradigmatic foundations are emerging in response to certain challenges that cannot be resolved using current mainstream approaches. The first proponents of theory-based evaluations were, in fact, reacting to black-box evaluations. Suchman (1967), Weiss (1972), and Chen and Rossi (1983) proposed that evaluations would have greater explanatory power if they included a well-founded logic model that detailed the action mechanisms of the intervention being evaluated. Chen (1990: 18) pointed out that "black box evaluations may provide a gross assessment of whether or not a program works but fail to identify the underlying mechanisms that generate the treatment effects, thus failing to pinpoint the deficiencies of the program for future program improvement and development" and that they are not sensitive to the influence of political and organizational contexts. The legitimacy of this type of evaluation was greatly advanced by Weiss (1972) and Suchman (1967), who noted that failure to find program effects could, when not attributable to faulty evaluation design, be due to either inadequate implementation or wrong theory (Bickman, 1987; Birckmayer and Weiss, 2000; Blinded for review; Chen, 2004; Weiss, 2007).

Constructing and analyzing program theory also appears to be an essential method for resolving the problems inherent in complex interventions (Dubois et al., 2012; Morell, 2010): How to deal with uncertainty created by interdependency among numerous actors who are constantly evolving and adapting? How to adapt to non-linear and sometimes unpredictable relationships? How to assess emergent and unanticipated outcomes resulting from relationships that are sometimes non-linear? (Morell, 2010; Shiell et al., 2008). Morell (2010) points out that program theory is crucial for evaluating this type of intervention because it helps reduce uncertainty. In 2012, Dubois and colleagues, in a special issue of the *Canadian Journal of Program Evaluation* on the evaluation of complex interventions (Houle et al., 2012) showed that evaluators employed program theory, in their evaluative practices, as a

core approach to better encompass the complexity of interventions (Dubois et al., 2012; Zimmerman et al., 2012). The explanatory power of a theory is helpful in anticipating the unexpected, understanding surprising phenomena, and reducing uncertainty (Morell, 2010). Likewise, the fifth generation underscores the explanatory power of contextual characteristics, implementation processes, and causal pathways to show, by identifying expected effects and impacts, how an intervention's activities and outputs led to outcomes. As the intervention unfolds, several implicit causal mechanisms result in the cumulative success or failure of the entire intervention or some of its components. Theory-based approaches to evaluation are used to shed light on these mechanisms that operate in open systems and are embedded in multiple social systems.

Our analysis suggests a new reading of theory-based approaches. This reading is not based on an exhaustive analysis of approaches, nor has it been validated by different theory designers. These two elements could bring significant information to the discussion.

Conclusion

Evaluation theoreticians often deplore the fact that new theories emerge without having been tested on the ground. The analysis of several theory-based approaches that have already been tested reveals the extent to which these approaches share similar foundations while being complementary in terms of the evaluation questions to which they can be applied. Together they form the basis of a new theory-based evaluation theory that could, in light of its historical evolution, mark the emergence of the 5th generation in evaluation.

Note

1. While there are some examples of theory-based economic evaluations, the current literature is not yet sufficiently well-structured that we can speak of a real approach. Indeed, several authors have chosen to use program theory in economic evaluation to report on effects that are not easily measurable. This is the case primarily for (social) return-on-investment analysis, which compares the costs of an intervention to its effects converted into dollars. The main challenge of social return-on-investment (SROI) analysis lies in how the evaluator represents and takes into account non-measurable, yet important, effects and maps the chain of effects and impacts. Several authors use program theory to capture all the impacts of an intervention (Banke-Thomas et al., 2015; Blinded for review; Nicholls et al., 2009, 2012; Tchouaket and Brousselle, 2013). With the recent demand for specific approaches in the public health field to report on diffuse, poorly delineated, and long-term effects (Drummond et al., 2007; Kelly et al., 2005) and the current fascination with SROI analyses (Eckermann et al., 2014; Edwards et al., 2013), we can expect to see an emerging standardization of economic evaluation methodologies along the lines of theory-based evaluations.

References

Alkin MC (ed.) (2004) *Evaluation Roots: Tracing Theorists' Views and Influences*. Thousand Oaks, CA: SAGE.

Alkin MC (2005) Utilization of evaluation. In: Mathison S (ed.) *Encyclopedia of Evaluation*. Thousand Oaks, CA: SAGE, 434–6.

Banke-Thomas AO, Madaj B, Charles A, et al. (2015) Social Return on Investment (SROI) methodology to account for value for money of public health interventions: A systematic review. *BMC Public Health* 15: 582.

Bhaskar R (2008) *A Realist Theory of Science*. New York: Routledge.

Bickman L (ed.) (1987) *Using Program Theory in Evaluation. New Directions for Evaluation, Vol*. 33. San Francisco, CA: Jossey-Bass.

Bickman L (ed.) (1990) *Advances in Program Theory. New Directions for Evaluation, Vol*. 47. San Francisco, CA: Jossey-Bass.

Biggs JS, Farrell L, Lawrence G, et al. (2014) A practical example of Contribution Analysis to a public health intervention. *Evaluation* 20(2): 214–29.

Birckmayer JD and Weiss CH (2000) Theory-based evaluation in practice: What do we learn? *Evaluation Review* 24: 407–31.

Brousselle A, Contandriopoulos D and Lemire M (2009) Using logic analysis to evaluate knowledge transfer initiatives: The case of the Research Collective on the Organization of Primary Care Services. *Evaluation* (London, England: 1995) 15(2): 165–83.

Brousselle A and Champagne F (2011) Program theory evaluation: Logic analysis. *Evaluation and Program Planning* 34: 69–78.

Brousselle A, Benmarhnia T and Benhadj L (2016) What are the benefits and risks of using return on investment to defend public health programs? *Preventive Medicine Reports* 3: 135–38. Available at: http://dx.doi.org/10.1016/j.pmedr.2015.11.015

Champagne F, Contandriopoulos A-P, Brousselle A, et al. (2011-2e édition) « L'évaluation dans le domaine de la santé : concepts et méthodes », dans Brouselle A., Champagne F., Contandriopoulos A-P., Hartz Z. (Eds) *L'Évaluation : Concepts et méthodes*, Presses de l'Université de Montréal : 49–70.

Champagne F, Contandriopoulos A-P and Tanon A (2011) Utiliser l'évaluation [Using evaluation]. In: Brousselle A, Champagne F, Contandriopoulos A-P, et al. (eds) (2011) *L'Évaluation : Concepts et méthodes*, 2nd edn. Montreal: Presses de l'Université de Montréal, 251–72.

Chen H-T (1990) *Theory-Driven Evaluations*. Newbury Park, CA: SAGE.

Chen H-T (2004) The roots of theory-driven evaluation: Current views and origins. In: Alkin M (ed.) *Evaluation Roots: Tracing Theorists' Views and Influences*. Thousand Oaks, CA: SAGE, 132–52.

Chen H-T and Rossi PH (1983) Evaluating with sense: The theory-driven approach. *Evaluation Review* 7(3): 283–302.

Contandriopoulos D, Brousselle A, Dubois C-A, et al. (2015) A process-based framework to guide nurse practitioners integration into primary healthcare teams: Results from a logic analysis. *BMC Health Services Research* 15: 78.

Contandriopoulos D, Brousselle A and Kêdoté NM (2008) Evaluating interventions aimed at promoting information utilization in organizations and systems. *Healthcare Policy* 4(1): 89–107.

Coryn CLS, Noakes LA, Westine CD, et al. (2011) A systematic review of theory-driven evaluation practice form 1990 to 2009. *American Journal of Evaluation 32*: 199–226.

Delahais T and Toulemonde J (2012) Applying contribution analysis: Lessons from five years of practice. *Evaluation* 18(3): 281–93.

Donaldson SI (2007) *Program Theory-Driven Evaluation Science: Strategies and Application*. Mahwah, NJ: Lawrence Erlbaum.

Drummond M, Weatherly H, Claxton K, et al. (2007) *Assessing the Challenges of Applying Standard Methods of Economic Evaluation to Public Health Interventions. Final Report*. London, UK:

Public Health Research Consortium. Available at: http://phrc.lshtm.ac.uk/papers/PHRC_D1-05_Final_Report.pdf

Dubois N, Lloyd S, Houle J, et al. (2012) Practice-based evaluation as a response to address intervention complexity. *Canadian Journal of Program Evaluation* 26(3): 105–13.

Dybdal L, Nielsen SB and Lemire S (2010) Contribution analysis applied: Reflections on scope and methodology. *Canadian Journal of Program Evaluation* 25(2): 29–57.

Easton G (2010) Critical realism in case study research. *Industrial Marketing Management* 39(1): 118–28.

Eckermann S, Dawber J, Yeatman H, et al. (2014) Evaluating return on investment in a school based health promotion and prevention program: The investment multiplier for the Stephanie Alexander Kitchen Garden National Program. *Social Science & Medicine* 114: 103–12.

Edwards RT, Charles JM and Lloyd-Williams H (2013) Public health economics: A systematic review of guidance for the economic evaluation of public health interventions and discussion of key methodological issues. *BMC Public Health* 13: 1001.

Funnell SC and Rogers PJ (2011) *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. San Francisco, CA: Jossey-Bass.

Groff R (2004) *Critical Realism, Post-positivism and the Possibility of Knowledge*. New York: Routledge.

Guba EG and Lincoln YS (1989) *Fourth Generation Evaluation*. Newbury Park, CA: SAGE.

Hewitt G, Sims S and Harris R (2012) The realist approach to evaluation research: An introduction. *International Journal of Therapy and Rehabilitation* 19(5): 250–9.

Houle J, Dubois N, Lloyd S, et al. (eds) (2012) L'évaluation des interventions complexes [Evaluation of complex interventions]. Special issue. *Canadian Journal of Program Evaluation* 26(3).

Jagosh J, Bush PL, Salsberg J, et al. (2015) A realist evaluation of community-based participatory research: Partnership synergy, trust building and related ripple effects. *BMC Public Health* 15: 725.

Jamal F, Fletcher A, Shackleton N, et al. (2015) The three stages of building and testing midlevel theories in a realist RCT: A theoretical and methodological case-example. *Trials* 16: 466.

Kelly MP, McDaid D, Ludbrook A, et al. (2005) Economic appraisal of public health interventions. *Briefing Paper, ISBN1-84279-431-0*. London, UK: NHS Health Development Agency.

Lemire ST, Nielsen SS and Dybdal L (2012) Making contribution analysis work: A practical framework for handling influencing factors and alternative explanations. *Evaluation* 18(3): 294–309.

Mackie JL (1965) Causes and conditions. *American Philosophical Quarterly* 2(4): 245–64.

Mayne J (2001) Addressing attribution through contribution analysis: Using performance measures sensibly. *Canadian Journal of Program Evaluation* 16(1): 1–24.

Mayne J (2008) Contribution analysis: An approach to exploring cause and effect. *The Institutional Learning and Change Initiative, Brief 16.* Rome, Italy. Available at: http://betterevaluation.org/sites/default/files/ILAC_Brief16_Contribution_Analysis.pdf

Mayne J (2011) Contribution analysis: Addressing cause and effect. In: Forss K, Marra M and Schwartz R (eds) *Evaluating the Complex: Attribution, Contribution and Beyond*. New Brunswick, NJ: Transaction Publishers, 53–96.

Mayne J (2012a) Contribution analysis: Coming of age? *Evaluation* 18(3): 270–80.

Mayne J (2012b) Making causal claims. Presentation at the International Program for Development Evaluation Training, *Building Skills to Evaluate Development Interventions*, 5–30 June, Ottawa, Canada. Available at: http://www.ipdet.org/page.aspx?pageId=videoJohnMayne2012

Mayne J (2015) Useful theory of change models. *The Canadian Journal of Program Evaluation* 32(2): 119–42.

Mehdipanah R, Manzano A, Borrell C, et al. (2015) Exploring complex causal pathways between urban renewal, health and health inequality using a theory-driven realist approach. *Social Science & Medicine* 124: 266–74.

Mingers J (2015) Helping business schools engage with real problems: The contribution of critical realism and systems thinking. *European Journal of Operational Research 242*: 316–31.

Morell JA (2010) *Evaluation in the Face of Uncertainty: Anticipating Surprise and Responding to the Inevitable.* New York: The Guilford Press.

Nicholls J, Lawlor E, Neitzert E, et al. (2009) *A Guide to Social Return on Investment*. London: The Cabinet Office, Office of the Third Sector. Available at: https://ccednet-rcdec.ca/files/ccednet/pdfs/2009-SROI_Guide_2009.pdf

Nicholls J, Lawlor E, Neitzert E, et al. (2012) *A Guide to Social Return on Investment*, 2nd edn. London: The Cabinet Office. Available at: http://socialvalueuk.org/what-is-sroi/the-sroi-guide

Patton MQ (1988) The evaluator's responsibility for utilization. *Evaluation Practice* 9(2): 5–24.

Patton MQ (2011) *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: Guilford Press.

Pawson R (2002) Evidence-based policy: The promise of 'realist synthesis'. *Evaluation* 8(3): 340–58.

Pawson R and Tilley N (1997) *Realistic Evaluation*. London: SAGE.

Pawson R and Tilley N (2004) Realist evaluation. Available at: www.communitymatters.com.au/RE_chapter.pdf

Pawson R and Tilley N (2005) Realistic evaluation. In: Mathison S (ed.) *Encyclopedia of Evaluation*. Thousand Oaks, CA: SAGE, 362–67.

Pawson R, Greenhalgh T, Harvey G, et al. (2004) Realist synthesis: An introduction. Submitted to the ESRC Research Methods Programme, Working Paper Series. Available at: http://betterevaluation.org/sites/default/files/RMPmethods2.pdf

Pawson R, Greenhalgh T, Harvey G, et al. (2005) Realist review – A new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy* 10(1): 21–34.

Quale A (2012) Radical constructivism on the Role of constructivism in mathematical epistemology. *Constructivism Foundation* 7(2): 104–11.

Ranmuthugala G, Cunningham FC, Plumb JJ, et al. (2011) A realist evaluation of the role of communities of practice in changing healthcare practice. *Implementation Science* 6(1): 49.

Rey L, Brousselle A and Dedobbeleer N (2012) Logic analysis: Testing program theory to better evaluate complex interventions. *Canadian Journal of Program Evaluation* 26(3): 61–89.

Ridde V, Robert É, Guichard A, et al. (2012) L'approche Realist à l'épreuve du réel de l'évaluation des programmes. *Canadian Journal of Program Evaluation* 26(3): 37–59.

Rogers P (2007) Theory-based evaluations: Reflections ten years on. *New Directions for Evaluation* 114: 63–7.

Rogers PJ, Hacsi TA, Petrosino A, et al. (eds) (2000) *Theory in Evaluation: Challenges and Opportunities*. *New Directions for Evaluation, Vol*. 87. San Francisco, CA: Jossey-Bass.

Salter K and Kothari A (2014) Using realist evaluation to open the black box of knowledge translation: A state-of-the-art review. *Implementation Science* 9(1): 115.

Sen A (1993) Positional objectivity. *Philosophy & Public Affairs* 22: 126–45.

Shadish W, Cook TD and Leviton LC (1991) *Foundations of Program Evaluation: Theories of Practice*. Thousand Oaks, CA: SAGE.

Shiell A, Hawe P and Gold L (2008) Complex interventions or complex systems? Implications for health economic evaluation. *BMJ* 336(7656): 1281–3.

Sridharan S and Nakaima A (2012) Towards an evidence base of theory-driven evaluations: Some questions for proponents of theory-driven evaluation. *Evaluation* 18(3): 378–95.

Stufflebeam DL and Coryn CL (2014) *Evaluation Theory, Models, and Applications*, 2nd edn. San Francisco, CA: Jossey-Bass.

Suchman EA (1967) *Evaluative Research: Principles and Practice in Public Service and Social Action Programs*. New York: Russell Sage Foundation.

Tchouaket E and Brousselle A (2013) Using the results of economic evaluations of public health interventions: Challenges and proposals. *Canadian Journal of Program Evaluation* 28(1): 42–66.

Tremblay M-C, Brousselle A, Richard L, et al. (2013) Defining, illustrating and reflecting on logic analysis with an example from a professional development program. *Evaluation and Program Planning* 40: 64–73.

Walsh D and Evans K (2014) Critical realism: An important theoretical perspective for midwifery research. *Midwifery* 30: e1–e6.

Weiss CH (1972) *Evaluation Research: Methods for Assessing Program Effectiveness*. Englewood Cliffs, NJ: Prentice-Hall.

Weiss CH (1988) Evaluation for decision: Is anybody there? Does anybody care? *Evaluation Practice* 9(3): 5–19.

Weiss CH (1997) How can theory-based evaluation make greater headway? *Evaluation Review* 21: 501–24.

Weiss CH (2007) Theory-based evaluation: Past, present and future. In: Mathison S. (ed.) *Enduring Issues in Evaluation: The 20th Anniversary of the Collaboration between NDE and AEA: New Directions for Evaluation 114*. San Francisco, CA: Jossey-Bass, 68–81.

Zimmerman B, Dubois N, Houle N, et al. (2012) How does complexity impact evaluation? An introduction to the special issue. *Canadian Journal of Program Evaluation* 26(3): v–xx.

Astrid Brousselle is a professor and director of the School of Public Administration at the University of Victoria. She received her Ph.D., in 2002, from the health administration department of the University of Montreal.

Jean Marie Buregeya, MD, MPH is a PhD candidate in Clinical Sciences Program at the Université de Sherbrooke and at the Research Centre Charles-LeMoyne - Saguenay -Lac-Saint-Jean on health care innovations in the Faculty of Medicine and Health Sciences, Sherbrooke University, Longueuil Campus.