CONTRIBUTION ANALYSIS APPLIED: REFLECTIONS ON SCOPE AND METHODOLOGY

Line Dybdal
Steffen Bohni Nielsen
Sebastian Lemire
Ramboll Management Consulting
Aarhus, Denmark

Abstract: This article investigates contribution analysis, an analytical tool invented by John Mayne, from both a theoretical and an operational perspective. Despite the substantial attention that contribution analysis has received, few studies appear to have applied it in practice. The article discusses the broadened scope of contribution analysis by analyzing its theoretical and methodological tenets, and examines its practical applicability in relation to two evaluations. The authors find that contribution analysis has much to offer the current theory-based evaluation landscape, but that further elaboration of the theoretical and methodological framework of contribution analysis is also needed.

Résumé : Cet article s'intéresse à l'analyse de contribution, telle que conçue par John Mayne, d'un point de vue à la fois théorique et pratique. Malgré une attention croissante il semble que peu d'études appliquent l'analyse de contribution. Cet article présente le champ élargi de l'analyse de contribution en analysant les tenants théoriques et méthodologiques, ainsi que leur application dans le cadre de 2 travaux d'évaluation en cours. Les auteurs concluent que l'analyse de contribution a beaucoup à offrir dans le contexte actuel de l'évaluation dite « théorique », mais une analyse plus poussée du cadre théorique et méthodologique de l'analyse de contribution est également nécessaire.

The attribution challenge is a key issue in the evaluation field, and, if anything, constitutes the boundary between the evaluation field and other fields such as auditing and monitoring. However, establishing causal links between interventions and discernible outcomes using textbook prescriptions for optimal research design is not always possible or even appropriate (Cook, Scriven, Coryn, & Evergreen, 2010). The Canadian evaluator John Mayne

Corresponding author: Line Dybdal, Ramboll Management Consulting, Olof Palmes Allé 20, DK 8200 Aarhus N, Denmark; <lind@r-m.com>

took note of this conundrum, and in a series of seminal papers and articles proposed a novel approach for addressing the question of attribution that he termed "contribution analysis" (CA) (Mayne, 1999, p. 3, 2001, 2008, 2011).

Over the years, Mayne's introduction of contribution analysis has attracted widespread attention from the global evaluation community (e.g., Hofmann, Roberts, Shoham, & Harvey, 2004; Lipski & Ignagni, 2001; Mackay, Horton, Dupleich, & Anderson, 2002; Newcomer & Schreirer, 2001). However, Mayne himself noted (2011) that, despite this substantial attention, there have been few published studies that systematically apply contribution analysis (Kotvojs & Shrimpton [2007] is an exception). The objective of this article accordingly is to better understand the analytical and practical application of contribution analysis, and hence how the apparent gap between the interest in CA and its actual application might be bridged. This will be achieved by assessing contribution analysis from both a theoretical and an operational perspective.

With this objective in mind, the article has been divided into five sections. In the first section we outline the methodological tenets of CA, paying particular attention to Mayne's recent application of CA as an alternative approach to meeting the attribution challenge in the context of evaluation. In the second section we situate CA in relation to the current debates taking place in the realm of theory-based evaluation (TBE). More specifically, we examine four limitations of theory-based evaluation that our experience has shown CA is capable of addressing. The third section reviews the practical application of CA on the basis of two evaluation studies. The fourth section presents a discussion concerning selected elaborations of CA that may render it more operational for evaluators. Fifth and finally, a brief conclusion summarizes the central points of this article.

OUTLINE OF CONTRIBUTION ANALYSIS

The term "contribution analysis" originates from financial management and business analysis, but has since been applied, with a different meaning, in the context of evaluation. In his inaugural work on the subject, John Mayne defined contribution analysis as follows:

> [a] specific analysis undertaken to provide information on the contribution of a program to the outcomes it is trying to influence. (Mayne, 1999, p. 6)

The initial objective of contribution analysis was to construct a performance story from existing data obtained from monitoring systems, but with limited collection of new data. At least initially, Mayne's ambition for CA did not appear to reflect or address the attribution question—in other words, the question of how causality between an intervention and a set of desired outcomes can be established. Rather, CA was proposed as a pragmatic approach to get the most out of performance measurement. In a later publication that elaborated on the operational design of CA, Mayne began to tackle the attribution challenge:

> We suggest undertaking a *contribution analysis* that would examine and present the best case possible—a credible performance story—for attribution with the available evidence. In essence, we are suggesting *theory-driven performance measurement.* (Mayne, 2001, p. 22; author's italics)

However, Mayne was still positioning CA in the realm of performance measurement.

Finally, in his more recent work, Mayne argues for CA to be employed in the evaluation of complex programs (2011). According to Mayne, CA is useful where it is impractical, inappropriate, or impossible to address the attribution question through an experimental evaluation design (2011). Thus, Mayne's objective is to provide an alternative and non-counterfactual way to address the attribution challenge in the context of evaluation.

The often heated debates surrounding the attribution challenge are well-documented elsewhere (see Cook et al., 2010; House, 2001; White, 2010), and there is therefore no need to dwell on those exchanges here. However, for the purpose of this article it is worth briefly noting two important developments that both motivate and frame Mayne's advancement of contribution analysis in the context of evaluation.

First, the notion of social causation has undergone significant developments during the decades of debate about the challenge of attribution (House, 2001). While the early positivist position viewed social causation in terms of linear cause-and-effect relationships, social causation is now recognized by most evaluators as being more akin to a complex and dynamic network of nonlinear and interacting cause-effect relations (House, 2001; Pawson & Tilley, 1997; White, 2010).

Second, there appears to be a growing acceptance, substantially informed by these underlying changes in the notion of social causation, that despite the many methodological strengths of counterfactual designs in providing certainty about causal inferences, determining causal links between interventions and discernible outcomes using textbook-prescribed optimal research design is not always possible or even appropriate (Cook et al., 2010).

Mayne shows that he is strongly aware of this evolution when he states that "in complex systems, experimenting with exogenous variables is not possible or not practical: the counterfactual case cannot be established" (2011, p. 4). Consequently, "contribution analysis does not involve a counterfactual-based argument, but rather builds a case for reasonably inferring causality" (Mayne, 2011, p. 6). As a consequence, Mayne goes on to argue, in these instances the evaluation question must be addressed by focusing on the intervention's contribution to a set of observed (positive or negative) outcomes (Mayne, 2011). Hence the challenge is not so much one of establishing, and providing proof, that the program caused the outcomes (i.e., demonstrating attribution), but rather of providing the best possible evidence that the intervention contributed to the outcomes of interest (i.e., demonstrating contribution), and hence of establishing the association between the program and its probable outcomes.

According to Mayne, one can infer probabilistic causation between an intervention and outcomes if five criteria concerning the embedded theory of change are met: (a) plausibility, (b) implementation according to plan, (c) evidentiary confirmation of key elements, (d) the taking into account of other influencing factors, and (e) the disproving of alternative explanations (2011). In this sense, contribution analysis provides an alternative approach to addressing the attribution challenge through its exploration of how a policy or program contributes to the observed results (Scottish Government, 2009).

As indicated above, CA has evolved over time, and we posit that three generations of contribution analysis can be discerned. The main developments are depicted in Table 1.

Scope

The main tenets remain the same throughout the three generations of CA: (a) elaborating the intervention's theory of change, (b) identifying key threats to the mechanisms of the theory of change, (c) identifying

Table 1
Overview of the Three Generations of Contribution Analysis

	First-generation CA	*Second-generation CA*	*Third-generation CA*
Article	Mayne (1999)	Mayne (2001)	Mayne (2011)
Scope	How monitoring data can be used to address the attribution question through CA (p. 5)	How monitoring data can be used to address the attribution question through CA (p. 5)	How CA can be used in evaluation and/or performance measurement (p. 1)
Problems addressed	Sketches elements in CA	Elaborates operational CA design	Elaborates operational CA design and considers implications in terms of causal claims
Operational approach	CA structured around nine elements	CA's nine elements consolidated into six steps	CA's six steps changed slightly. New Step 1 and Steps 3 & 4 consolidated as new Step 4. Step 7 added
Level of detail for each step	Outline guidelines	General guidelines	Elaborate guidelines
On attribution	Acknowledges experimental design most appropriate for addressing attribution (p. 5)	Acknowledges experimental design most appropriate for addressing attribution (pp. 4–5) Argues that "plausible association" can be inferred (p. 8) Contribution can be inferred when four conditions are met (pp. 21–22)	Argues that other approaches to attribution can be valid (pp. 5–6) Argues that probabilistic causation can be inferred (p. 7) Contribution can be inferred when four conditions are met (p. 7)
On magnitude of contribution	None offered	None offered	Based on evidence available and causal links demonstrated three levels of contribution can be determined (pp. 25–26).
Methodology	Advocates mix of quantitative and qualitative methods. Use of existing data and additional data collected (p. 5)	Same (p. 18)	Same (pp. 17–21)
Epistemology	None offered	None offered	None offered
Ontology	None offered	None offered	None offered

other contributing factors, and (d) testing the principal competing explanations. The most significant development in CA is the broadening of its scope from the context of performance measurement to include evaluation studies. In his most recent work on CA, Mayne (2011) suggests that this approach can also be applied to the evaluation of complex programs in order to assess their contribution to a set of outcomes. He goes on to argue that three basic kinds of contribution story can be told, depending on the relative strength of the evidence.

A *minimalist contribution analysis* can be posited when a theory of change was developed and the expected outputs were delivered. The assertion of contribution is based on "the inherent strength of the postulated theory of change and the fact that the expected outputs were observed" (Mayne, 2011, p. 25).

A *contribution analysis of direct influence* can be posited when a theory of change was developed, expected outputs occurred, immediate results were observed, and the evidence suggests that the program was instrumental in creating those results in the presence of other influencing factors.

A *contribution analysis of indirect influence* can be posited when

> [i]t would measure the intermediate and final outcomes (or some of them) and gather evidence that the assumptions (or some of them) in the theory of change in the areas of indirect influence were borne out. Statements of contribution at this level would attempt to provide factual evidence for at least the key parts of the whole postulated theory of change. (Mayne, 2011, pp. 25–26)

Operational Approach

Mayne initially identified nine elements in contribution analysis (1999). He subsequently consolidated these into six steps (2001), which he slightly modified with the addition of a seventh step for the evaluation of complex interventions (2011).

In his conception of the theory of change, Mayne increasingly stresses not only the implementation theory, but also the program theory and the risks associated with its assumptions (1999, 2001). In his later work Mayne suggests that the results chain, the underlying assumptions and risks, and the contextual and other influencing factors must

all be included in what he terms an "embedded theory of change," as well as stating what criteria must be met in order to construct such a theory (2011, pp. 16–17). This elaboration is only one element of an increasing level of detail and guidance for each step of a contribution analysis.

Importantly, the first four steps of CA are to be carried out in what is usually referred to as the structuring phase of an evaluation (e.g., Nielsen & Ejler, 2008). The remaining steps are to be executed in the data collection and analysis phases of an evaluation.

Methodology

Kotvojs and Shrimpton note some misconceptions as to what methods must be applied in CA (2007; Kotvojs, 2006). Suffice it to say that in his own work Mayne remains insistent that a mix of quantitative and qualitative methods can be used, and that it is feasible to both use existing data and collect additional primary data. The weighting given to the respective data type naturally depends on the scope of the evaluation (Mayne, 1999, 2001, 2011).

Attribution

An interesting development in Mayne's perception of, and ambition for, CA as a means of addressing the challenge of attribution is evident. In his initial works he appears to summarily acknowledge that the attribution question is best addressed through experimental design, but does not go into details (Mayne, 1999, 2001). He initially focuses on reducing uncertainty through evaluative inquiry by arguing that one can infer "plausible association" (Mayne, 1999, pp. 5–7). What is entailed in inferring plausible association is further elaborated in a later work (Mayne, 2001), and he ultimately places CA in the debate concerning social causation by stating how a causal inference can be made. This evolution is depicted in Table 2.

The question of whether CA is a credible alternative to traditional counterfactual designs lies beyond the scope of this article. However, when Mayne asserts that probabilistic causation can be established by testing the theory of change in an evaluation, he inserts CA into the current debates concerning theory-based evaluation (Mayne, 2009), and it must therefore be assessed in this context. We will explore this issue in more detail below.

Table 2
On CA's Understanding of Attribution

	First-generation CA	*Second-generation CA*	*Third-generation CA*
Article	Mayne (1999)	Mayne (2001)	Mayne (2011)
On attribution	Acknowledges experimental design most appropriate to address attribution (p. 5)	Acknowledges experimental design most appropriate for addressing attribution (pp. 4–5) The analysis must demonstrate "plausible association" between intervention and outcomes (p. 8)	Argues that other approaches to attribution can be valid (pp. 5–6) Argues for probabilistic causation where one can infer a causal relation between intervention and outcomes with reasonable confidence (p. 7)
On magnitude of contribution	None offered	None offered	Based on evidence available and causal links demonstrated: 1. Minimalist contribution analysis 2. Contribution analysis of direct influence 3. Contribution analysis of indirect influence (pp. 25–26)

CONTRIBUTION ANALYSIS AND THEORY-BASED EVALUATION

In this section we outline the challenges identified in the theory-based evaluation debates and how CA potentially contributes to these challenges.

The initial methodological debates concerning theory-based evaluation primarily revolve around such issues as (a) the need for theory-based evaluation, (b) the notion of theory, and (c) the construction of theory (Bickman, 1987, 2000; Chen, 1990, 1994; Chen & Rossi, 1983, 1987; Schreirer, 1987).

TBE was initially introduced by means of several different approaches to evaluation that could be used as either an alternative or a supplement to counterfactual designs for addressing the attribution question when methodological, ethical, or practical constraints had to be taken into account. The objective was to gain insight into the black box of causal relationships by carefully modelling the structural relationships that existed between relevant variables and intervening processes (Chen & Rossi, 1987). Chen and Rossi (1983) argued that

using program theory during evaluation would significantly improve the impact of the latter, and asserted the need to develop a plausible and defensible program theory before embarking on an evaluation.

Accordingly, the construction of a sound program theory is the foundation on which the evaluator would have to test the causal relationships. Therefore, a central task is to identify which causal links to test (Chen & Rossi, 1983). Chen later elaborated on how the theory should be used, and argued that the program theory should be the conceptual framework for designing the subsequent phases of the evaluation. In this way, the use of program theory not only strengthened the validity of the study, but enhanced communication between stakeholders and evaluators in relation to the design of useful evaluations (Chen, 1990). In the early stages, however, much writing on TBE was based on theoretical work rather than on empirical experience based on the actual use of TBE in connection with the conducting of evaluation studies, and only a minor element of the theoretical work was concerned with empirically testing the theories using different techniques (Mark, 1990; Marquart, 1990; Smith, 1990).

During the course of subsequent debates, more focus has been directed toward defining TBE compared with other approaches to evaluation, and also toward developing minimum standards for conducting TBE. There is widespread recognition that TBE is conceptually and operationally based on an explicit theory that describes how a program produces certain outcomes, and that an evaluation is at least partly guided by this theory (Rogers et al., 2000). Coryn, Noakes, Westine, and Schröter further define TBE as "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" (2011, p. 3). Consequently, TBE can be thought of as being based on five core principles: (a) theory formulation; (b) theory-guide question formulation; (c) theory-guide evaluation design, planning, and execution; (d) theory-guide construct measurement; and (e) causal description and causal explanation (Coryn et al., 2011).

The Practical Limitations of TBE

Although TBE has received a lot of academic attention, less attention has been given to the practical application of TBE approaches. A study conducted by Patricia Rogers identified few full-blown examples of TBE, but "many interesting variations of PTE [Program Theory

Evaluation] in practice and much to recommend it" (Rogers et al., 2000, p. 5). Since then, Donaldson (2010) has found a number of evaluation studies that apply a TBE approach, and reflects on the use of the approach in relation to (a) developing program impact theory, (b) formulating and prioritizing evaluation questions, and (c) answering evaluation questions. In their reviews of practical TBE approaches, Carol Weiss (1997) and Patricia Rogers (Rogers, 2007; Rogers et al., 2000) highlight three particular pitfalls when conducting a TBE, and Peter Dahler-Larsen (2001) and Jane Davidson (2000) both highlight a fourth pitfall.

First, TBEs are often based on an implementation theory, which focuses on how the program is carried out, rather than on a program theory, which focuses on how a program is supposed to work (Weiss, 1997). Consequently, the evaluation will assess whether the program is being conducted as planned rather than assessing the (causal) mechanisms. This is not a problem when the aim is to gather knowledge on the implementation process. However, it is a significant problem when seeking to determine attribution. A systematic review of the practical implications of TBE shows that all studies enact a type of theory formulation (Coryn et al., 2011, p. 14), but does not state whether this theory is an implementation theory or a program theory.

Second, TBEs are often based on excessively simplistic or partial theories that have been developed primarily on the basis of the practitioners' assumptions (Rogers, 2007; Weiss, 1997). A simplistic program theory makes it harder "to trace the micro-steps of process all along the pathways that lead to program effects" (Weiss, 1997, p. 78), which also makes it harder to assert a reasonable claim to contribution. A partial program theory fails to use social theory and existing evidence that might be able to inform the identification of alternative explanations. In other words, a simplistic or partial program theory leads to the generation of simplistic or partial knowledge and claims to contribution. The recent review by Coryn et al. indicates that existing theory and research is increasingly being used in theory formulation (Coryn et al., 2011, p. 14).

Third, the theory is often not used for guiding the evaluation, and when it is, it is primarily concerned with answering the basic question of whether every component of the program theory has been implemented as planned (Davidson, 2000; Rogers, 2007; Weiss, 1997). The consequence is that the assumptions behind the theory's causal mechanisms remain untested. In later discussions this is found to be crucial to a TBE approach (Coryn et al., 2011; Donaldson, 2010).

Nevertheless, Coryn et al. found that 76% of the TBE studies they reviewed made use of theory to formulate evaluation questions, while only 22% used it to prioritize evaluation questions. The review also shows that approximately 40% of the studies made use of theory-guided measurement construction, and 50–60% used theory to explain cause-and-effect associations between theoretical constructs (Coryn et al., 2011, p. 14).

Fourth, minimal attention has been paid to alternative explanations or alternative theories of how the program is functioning (Davidson, 2000; Larsen, 2001). This implies that the overall findings of the program's contribution to the identified outcomes might easily be challenged, whereas a careful consideration of possible alternative explanations would strengthen the overall contribution story.

As we noted above, broadening the scope of CA to include the evaluation of programs requires Mayne to clarify the connection with other modes of theory-based evaluation. We posit that CA addresses all four of the above-mentioned challenges to TBE.

First, as far as the question of implementation versus program theory is concerned, CA can be regarded as a promising approach for expanding the use of program theories, because in his concept of the "embedded theory of change" Mayne (2011) is increasingly addressing not only the results chain, but also the underlying assumptions and risks, influencing factors, and principal competing explanations. Consequently, CA can be used to systematically uncover and specify intervening variables, and to discriminate between theory failure and implementation failure.

Second, Mayne's elaboration of the necessary components of an embedded theory also has the effect of improving the quality of the theory used. Focusing on the nexus of causal mechanisms underlying the program to be evaluated is crucial to the theoretical and practical structure of CA. Additionally, Mayne's attention to elaborating the influencing (program-internal and -external) factors and competing explanations also has the effect of making the theory underlying the evaluative inquiry more robust.

Third, as far as the question of the use of theory to guide the evaluation is concerned, CA represents a methodology that permits the theory of change to be integral to the analytical strategy. As has been mentioned above, a common criticism of TBE is the absence of systematic testing of the program theory's most important linkages

(Davidson, 2000; Rogers, 2007; Weiss, 1997). When Mayne's approach is applied systematically, both existing and novel evidence is collated and assessed in order to identify and test not only the program theory's internal causal linkages, but also the influence of other contributing factors or competing explanations.

Fourth and finally, as far as testing for alternative explanations is concerned, CA focuses explicitly on identifying and testing alternative or competing explanations.

In our view, CA represents a promising direction for TBE because it tackles a number of its challenges. So far, Mayne has only tangentially addressed this area (2009).

Ultimately, the applicability and utility of CA ought to be tested in the setting of a real-world evaluation. So in the following section we investigate the practical application of CA in order to assess its strengths and weaknesses.

THE APPLICATION OF CONTRIBUTION ANALYSIS

The basis for assessing the practical application of contribution analysis is the evaluation of two programs in which CA governed both the design and the analytical strategy. Both cases, which concern at-risk families and children, arise from a Danish social policy context. The first program, known as The Incredible Years (TIY), involves an evidence-based method for reducing the severity and prevalence of behavioural problems in children. The second program, known as the Parental Injunction Program (PIP), involves withholding the transfer payments made to parents if they fail to comply with an order imposed by social services, such as an order made to ensure that their child attends school.

The cases chosen reflect the application of CA to two quite different evaluation designs whose components range from time-series design to post-festum measurement, and which therefore also create different conditions for testing the applicability of CA. The timing of the evaluations also differed substantially. In the case of TIY, the evaluation commenced near-simultaneously with the program. In the case of PIP, the evaluation was retrospective.

Both evaluations were led by the same evaluator, but involved different evaluation teams. The appraisal of the applicability of CA that is

described below emerged from internal team discussions plus discussions with the quality reviewer both during and after the studies. The reviews focused on the applicability and utility of the CA in relation to the evaluations.

An overview of the cases is presented in Table 3.

Table 3
Cases of Applied Contribution Analysis

Name of study	Evaluation of the Parental Injunction Program (PIP)	Evaluation of the program titled The Incredible Years (TIY)
Commissioner	Danish Board of Social Services	Danish Board of Social Services
Evaluator	Ramboll Management Consulting (2007)	Ramboll Management Consulting (2010, 2011)
Sample	Case managers and parents in Danish municipalities	Administrators, program practitioners and stakeholders in municipalities
Context	Ex-post evaluation study of the PIP	Formative evaluation of TIY for at-risk families program
Level	Program	Program
Evaluation questions	The effects of the legislation Conditions affecting the target group that influence usage and effectiveness Which conditions in the municipality increase usage and effectiveness Circumstances that explain negative side effects	The effects of the program Fidelity of program implementation Program satisfaction among parents and program practitioners
Methodology	Case studies using interviews, documents, and surveys	Time-series analysis without comparison group Desk research Case studies
Limitations	Limited treatment population No baseline and endline data for effect indicators	Limited treatment population
Main findings	The parental injunction was effective in 10 of 24 cases and proved ineffective in 7 cases Local implementation influenced use of measure. Knowledge, guidelines, and attitude influenced implementation	Using validated scales a significant positive post-treatment development was noted for children. There was no significant development for parents Strong and visible leadership, resources for internal training, and regular supervision influenced implementation.

In the following, we outline the actions that John Mayne has prescribed for each step of the contribution analysis. Subsequently, we review its application in the two cases.

Step 1: Setting Out the Cause-Effect Issue To Be Addressed

Contribution analysis guideline. The first step in Mayne's operational approach to contribution analysis is to set out the cause-effect issue to be addressed. According to Mayne, this consists in acknowledging the attribution problem, which is to say recognizing that the attribution challenge needs to be addressed. An important step in acknowledging the attribution challenge is to scope the attribution problem, which involves determining both the specific cause-effect question being addressed and the level of confidence needed for answering the question.

Application. In both cases the evaluation was commissioned through a public procurement process, which meant that the terms of reference specified evaluation questions centring on cause-effect relationships. In the case of the TIY program, the timing and resources allowed for a randomized controlled trial, but the TIY was at an early stage of implementation and the commissioner and evaluator decided that it was more important to focus on implementation-related issues than to achieve the highest possible level of confidence (see also Hunter, 2006). For the PIP program, no standardized baseline or endline data were available. In other words, only ex-post measurements could be established.

In both cases, further scoping of the evaluation question was needed, and the procurement procedures involved extra iterations in connection with the issues described in Step 1. Consequently there was quite extensive (re)scoping of the cause-effect questions during Steps 2–4.

Concerning the practical application of CA in the two cases, we found that establishing the level of confidence often boils down to balancing methodologically optimal design against the practical limitations (i.e., terms of reference, availability of data, timing, and resources). The issue is further complicated by the fact that much evaluation work is commissioned through a public procurement process that restricts dialogue and scoping. Accordingly, the scoping of the evaluation questions and the decision about the most feasible methodology for answering them is often negotiated on the basis of an evaluation commissioner's terms of reference plus a proposal from the evaluator. A key issue in the two cases was to develop a clear and shared understanding of the basic contribution questions being asked, and to ensure that the commissioner understood the methodological implications of focusing on contribution rather than attribution. This was discussed in methodology notes approved by the commissioner by the end of Step 4.

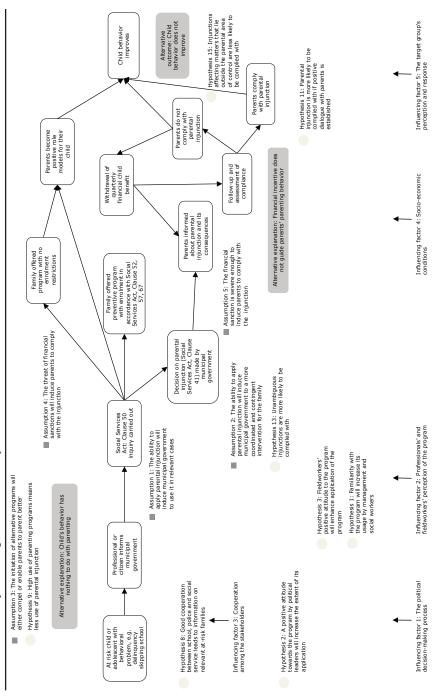Step 2: Developing the Postulated Theory of Change, and the Risks to the Theory

Contribution analysis guideline. According to Mayne (2011), the key process in contribution analysis involves using the embedded theory of change to drive the analytical strategy.

Influencing factors are contextual factors that either propel or inhibit the mechanisms driving the intended change, and they correspondingly either strengthen or weaken these mechanisms. Conversely, alternative explanations are (typically a complex of) mechanisms that are separate from the mechanisms expected to drive the intended changes, and they correspondingly challenge the overall theory.

The CA guidelines are fairly prescriptive concerning the composition of the embedded theory of change. Likewise, they are fairly precise about what composes a good theory of change. Citing Connell and Kubisch (1998), Mayne argues that a good theory of change is (a) plausible, (b) agreed, (c) embedded, and (d) testable. However, although the guidelines are less prescriptive about how to develop the embedded theory of change, Mayne does refer elsewhere to methodologies that focus on this area.

Application. For each program, the theory of change underlying the program was elicited and influencing factors were identified. Data concerning input, activities, output, and outcomes (still to be collected) were identified. In the case of TIY, the evaluator sought to identify influencing factors by surveying published evaluation studies concerning TIY in a selected number of clearing houses with (a) the same target population, and (b) the same, or very similar, intended outcomes. For the PIP case, the evaluator drew on research concerning programs in other countries that operate similar mechanisms of financial sanction. Both cases also drew on implementation research as a means of identifying influencing factors within or outside the program that drive or inhibit program performance (e.g., Fixsen, Naoom, Blasé, Friedman, & Wallace, 2005; Winther, 2001). In both cases, the evaluation teams identified positive and negative side effects, but struggled to construct alternative explanations in operational terms (see Figure 1). The evaluation teams found that the guidelines for identifying influencing factors and alternative explanations, and sometimes for distinguishing between them, were too vague to be operationally applicable. This is notable when taking into account that the embeddedness of the theory of change is one of the key arguments

Figure 1
Embedded Theory of Change for Parental Injunction

- Assumption 3: The initiation of alternative programs will either compel or enable parents to parent better
- Hypothesis 9: High use of parenting programs means less use of parental injunction

Alternative explanation: Child's behavior has nothing to do with parenting

At risk child or adolescent with behavioral problem, e.g. delinquency skipping school

Professional or citizen informs municipal government

Social Services Act: Clause 50 inquiry carried out

Family offered program with no enrolment restrictions

Parents become positive role models for their child

Child behavior improves

Alternative outcome: Child behavior does not improve

- Hypothesis 15: Injunctions affecting matters that lie outside the parental area of control are less likely to be complied with

Family offered preventive program with enrolment in accordance with Social Services Act, Clause 52, 57, 67

Decision on parental injunction, (Social Services Act, Clause 41) made by municipal government

Withdrawal of quarterly financial child benefit

Parents do not comply with parental injunction

Parents comply with parental injunction

Parents informed about parental injunction and its consequences

Follow-up and assessment of compliance

- Hypothesis 11: Parental injunction is more likely to be complied with if positive dialogue with parents is established

Influencing factor 5: The target group's perception and response

- Assumption 4: The threat of financial sanctions will induce parents to comply with the injunction

- Assumption 1: The ability to apply parental injunction will induce municipal government to use it in relevant cases

- Assumption 5: The financial sanction is severe enough to induce parents to comply with the injunction

Alternative explanation: Financial incentive does not guide parents' parenting behavior

- Assumption 2: The ability to apply parental injunction will induce municipal government to a more coordinated and contingent intervention for the family

- Hypothesis 13: Unambiguous injunctions are more likely to be complied with

Influencing factor 4: Socio-economic conditions

- Hypothesis 8: Good cooperation between school, police and social service leads to information on relevant at-risk families

Influencing factor 3: Cooperation among the stakeholders

- Hypothesis 2: A positive attitude towards the program by political leaders will increase the extent of its application

- Hypothesis 3: Fieldworkers' positive attitude to the program will enhance application of the program

- Hypothesis 1: Familiarity with the program will increase its usage by management and social workers

Influencing factor 2: Professionals' and fieldworkers' perception of the program

Influencing factor 1: The political decision-making process

for using CA compared with other theory-based approaches. In the case of TIY, the evaluation team attempted to identify influencing factors and alternative explanations by partially reversing Steps 2 and 3. They did this by using a literature review as the foundation for developing hypotheses concerning influencing factors and alternative explanations.

Step 3: Gathering the Existing Evidence on the Theory of Change

Contribution analysis guideline. According to Mayne, once a postulated theory of change has been assembled and risks have been assessed, evidence must be gathered regarding the component parts of the embedded theory of change. He outlines two elements in this step: (a) assessing the logical robustness of the theory of change, and (b) gathering evidence regarding results, assumptions, and other influencing factors (Mayne, 2011).

Mayne provides a practical guide to resources, plus criteria that can be useful for assessing the robustness of the theory of change (see Step 2).

According to Mayne, assessing the robustness of the theory is very similar to constructing a logic of probability, and is comparable to the way a lawyer builds a prosecution case. Evidence that key outputs were delivered, process standards were maintained, and the immediate, intermediate, and end outcomes were obtained all contributes to strengthening the case for the theory's robustness. Likewise, circumstantial evidence can be introduced, such as a precedent for a particular mechanism affecting a similar target population in a similar context.

Other influencing factors, such as program-internal and program-external inhibitors and drivers, must be mapped and their potential influence assessed. Again, the evidence derived from meta-evaluation and logical inference go hand in hand.

Application. As has been mentioned above, in both cases the evaluation teams arrived at two findings. First, Steps 2 and 3 effectively coincide and could beneficially be merged. Second, the evaluation team found that the actions suggested by Mayne in Step 3 had to be reversed. Assessing the robustness of the theory of change requires having the available evidence at hand. The relevant issues were therefore addressed in the previous step.

Step 4: Assembling and Assessing the Contribution Story, and the
Challenges to It

> *Contribution analysis guideline.* Mayne prescribes the following ac-
> tions in this step: (a) set out the contribution story, and (b) assess
> the strengths and weaknesses in the postulated theory of change in
> light of the available evidence and the relevance of other influencing
> factors.

> This implies that a preliminary outline and assessment of the con-
> tribution story must be based on existing evidence: in other words,
> it must be based on an interim reporting on the plausibility of the
> contribution story that goes beyond simply applying the embedded
> theory of change in order to understand the program.

> *Application.* In both cases, this step was used for more than an in-
> terim reporting on the plausibility of the contribution story. In both
> cases, the evaluators used this step to translate the embedded theory
> of change into an operational framework for collecting, analyzing,
> and reporting data using an evaluation matrix (Imas & Rist, 2009).

> First, the evaluation team developed a gross list of hypotheses to be
> tested through the subsequent collection of new evidence. The hypoth-
> eses were constructed on the basis of the work done in the preceding
> phases, and they addressed (a) risks to the main assumptions, (b)
> other influencing factors, and (c) alternative explanations. In order to
> make the theory of change guide the subsequent analysis, these ele-
> ments were consolidated into the embedded theory of change. Figure
> 1 depicts the theory of change that underlies the PIP. It illustrates the
> relationships between the hypotheses, the program, the influencing
> factors, and alternative explanations.

> Second, in the case of TIY the evaluation team prioritized a net list
> of hypotheses from which the most important candidates for testing
> had to be selected with a view to developing a sound foundation for
> the contribution analysis. The relative importance and strengths of
> the hypotheses were weighed against (a) the strength of evidence, (b)
> their proximity to the evaluation questions, and (c) their relevance
> to the interests of clients and stakeholders. In all, about a dozen hy-
> potheses were formulated for each program (see Figure 1). For both
> programs, the hypotheses were approved by the commissioner and
> steering group.

> Third, the prioritized hypotheses were operationalized as specific
> research questions, and strategies for collecting new evidence and

testing the hypotheses were developed. As far as assembling and assessing the contribution story is concerned, Mayne's model leaves evaluators with little more than examples of relevant questions to be asked: in practice, the evaluation team found that this step required further specification regarding actions and scoping. In each case, the hypotheses and evaluation methodology were consolidated in an interim report that was approved by the evaluation steering group.

Step 5: Seeking Out Additional Evidence

Contribution analysis guideline. According to Mayne, Step 5 is to seek out additional evidence. He outlines three elements in this step: (a) determine what kind of additional evidence is needed to enhance the credibility of the contribution story, (b) refine the theory of change, and (c) gather new evidence.

Mayne states that the results from assessing the contribution story (Step 4) will determine what further evidence base is needed for strengthening the contribution story.

At this point it will be apparent which are the most important evaluation questions, and which hypotheses should be prioritized. They have become part of an analytical strategy (see Step 4) that is to be implemented during this step.

Application. In both programs, the CA approach was used to produce answers to the evaluation questions concerning causation. In each program the evaluator systematically collected data that would test the hypotheses of the embedded theory of change regarding causal mechanisms and other influencing factors. Likewise, this was done using a mixed-method (drawing on quantitative and qualitative data) and multi-informant perspective (involving clients, social workers, middle and senior management, and experts). However, the differences in the evaluation designs implied a difference in the sophistication of the statistical analysis and the robustness of the findings. In the case of the TIY program, stronger inferences could be drawn concerning cause-effect relations, and therefore a stronger contribution story could be constructed.

In both cases, the evaluation team found that the CA approach had strengthened the analytical strategy and had focused the data collection efforts toward answering the aforementioned evaluation questions. However, the strength of the probable association between

a program and its outcomes ultimately hinges on the nature of the research design in which the CA is embedded.

Step 6: Revising and Strengthening the Contribution Story

Contribution analysis guideline. After the available evidence has been collected, collated, and analyzed, Mayne's guidelines prescribe the following: (a) construct a more credible contribution story, (b) reassess its strengths and weaknesses, and (c) revisit Step 5.

In these guidelines it is evident that Mayne's model is based on the notion of "probabilistic causation," where a given contribution is not proven but is indicated through an iterative process of continuously testing the contribution story and the risks to it (Mayne, 2011, p. 23). In other words, after Step 6 you must return to Step 5 to seek out additional evidence in order to revise and strengthen the contribution story even further.

Application. Contrary to the instructions contained in the methodology notes, the final evaluation reports give little attention to evidence supporting alternative explanations (Ramboll Management Consulting, 2007, 2011). Two concurrent explanations were found in each case: (a) the evaluation teams found it difficult to delineate and operationalize this concept, and (b) client expectations concerning the evaluation report meant that a low priority was given to this issue.

Both programs contractually specified final evaluation report deliverables. At first glance, Mayne's iterative approach of continuously establishing a more robust contribution story is appealing and useful in the context of performance management. However, the iterative process is somewhat at odds with its application in a one-off evaluation study. In both cases, time and resource limitations imposed obvious constraints on an iterative process; even more importantly, client expectations were also an obstacle to iterative and continuous testing leading to a more credible contribution story. In the case of TIY, the report was needed as a deliverable in order to determine what measures were required for replicating the program. In the case of PIP, revisions to the program's legal basis were imminent. In other words, an instrumental use of the evaluation reports by a specific date was foreseen.

These instances suggest that Mayne's ideas may not fare well in the context of genre conventions, contractual obligations, and client expectations for evaluations.

One potential way to bridge the gap between the intended and actual use of CA may be to make recommendations proposing a monitoring and evaluation strategy that could strengthen the program's future evidence base.

In the two cases investigated, the evaluation teams that had worked on the two finalized evaluations found that the systematic hypothesis testing on (a) assumptions and risks, (b) other influencing factors, and (c) alternative explanations significantly aided the foci comprising the analysis, judgement, conclusions, and recommendations regarding program errors, implementation errors, and the drivers and inhibitors that influence program outputs or desired results.

DISCUSSION

The expansion of CA into the sphere of program evaluation implies that CA will also have an impact on other genre conventions and expectations. Although the methodological gap between performance measurement and program evaluation may be exaggerated, differences do exist (see Nielsen & Ejler, 2008, for a discussion of this point). In the preceding pages we have surveyed the theoretical discussions concerning attribution and theory-based evaluations that include CA. Equally, we have analyzed the tenets of CA and its practical application in program evaluation. The motivation behind our effort is our experience of CA as a potentially strong and credible approach for making causal claims in the context of evaluation. However, we are also well aware that

> [a] stumbling block to using contribution analysis may be that this type of analysis remains seen, especially in some of the evaluation profession, as a very weak substitute for the more traditional experimental approaches to assessing causality. (Mayne, 2011, p. 1)

That being the case, there is still further work to be done. In proposing CA as a methodologically sound approach, we first discuss to what extent CA, when it is applied, successfully addresses the commonly posed challenges in theory-based evaluation that have been presented above.

Second, we assess CA from an operational and utility-related perspective, and assess what elements in the CA methodology will need to be adapted when evaluation studies are being conducted.

Insights on Application

First, it has often been asserted that TBE tends to focus on implementation theory rather than program theory. The actual cases showed that the concept of an embedded theory of change is tenable when both implementation theory and program theory are made fundamental components in understanding and assessing an intervention.

Second, much TBE work has been criticized for embodying excessively simplistic theories of change. As has been discussed above, the cases demonstrated that addressing all the elements in the embedded theory of change will deal with this weakness. However, in practice the cases also revealed that the methodology for identifying influencing factors and competing explanations needs to be further elaborated in order to be able to inform the evaluators' application of CA.

Third, the absence of a theory to guide the evaluation design and analytical strategy has drawn criticism. In the cases investigated, CA proved to be useful for guiding and focusing the designs and analytical strategies applied. In both cases, the CA guidelines needed to be made more concrete in relation to the development of sets of hypotheses for guiding the analytical strategy in subsequent work.

Fourth, the absence of a test for alternative explanations and for accounting for external factors has often been cited as a weakness of TBE. One intended strength of CA is that it focuses explicitly on identifying and accounting for alternative explanations and external factors. As was noted earlier, Mayne asserts that one can only infer probabilistic causation between an intervention and outcomes if the embedded theory of change accounts for other influencing factors and disproves alternative explanations (2011). However, while Mayne consistently emphasizes the need for identifying and accounting for other influencing factors, and for testing alternative explanations, in practice he is not very prescriptive about how to do this from a methodological perspective.

The two instances of the practical application of CA under discussion also reveal problems in operationalizing these dimensions, especially when it comes to alternative explanations. These cases suggest that the challenges may involve both the methodology (the *how to do it* challenge) and the presentation of evidence to clients and stakeholders (the *how to present it* challenge).

Insights on Utility

There is general acceptance that an evaluation progresses through different phases that entail designing, structuring, collecting data for, and analyzing the evaluation, not to mention judging and utilizing the evaluative inquiry and its findings.

Our first finding was that CA's stepwise approach may beneficially be linked to these phases in order to highlight the appropriate points for carrying out the tasks that each step entails.

Second, the contribution story may have to be constructed in the format of an evaluation report, at least for more cautious clients and stakeholders. CA needs to take these genre conventions into account.

Third, the iterative process for refining the contribution story was perceived as being unlikely to be applicable in the context of many commissioned evaluations.

Table 4 summarizes the observations made in the cases investigated concerning CA's applicability and utility. These observations may

Table 4
Proposed Adaptations to Contribution Analysis Methodology

Steps in contribution analysis	Proposed adaptations	To be executed in phase
Step 1: Set out the cause-effect issue to be addressed	Identify influencing factors at a later stage	Ideally in the design or structuring phase
Step 2: Develop the postulated theory of change and the risks to it	Clarify the method for identifying alternative explanations Clarify how to identify influencing factors In practice, Steps 2 and 3 occur almost simultaneously. Consider consolidating them into a single step	Structuring
Step 3: Gather the existing evidence for the theory of change	Reverse actions in this step. Gather evidence before assessing strengths and weaknesses Use implementation studies Generate gross list of hypotheses	Structuring
Step 4: Assemble and assess the contribution story plus the challenges to it	Develop deep program understanding Prioritize net list of hypotheses	Structuring
Step 5: Seek out additional evidence	Elaborate how alternative explanations can be accounted for in the reporting	Data collection
Step 6: Revise and strengthen the contribution story	Adapt to the genre conventions of evaluation studies	Reporting

further inform the elaboration of CA as a distinct and operational approach to theory-based evaluation.

CONCLUSION

In this article the authors have assessed contribution analysis from both a theoretical and a practical perspective. This has been done in the context of CA's ability to address the challenges experienced with TBE approaches, as well as in terms of its practical applicability.

When we reviewed CA's theoretical and methodological tenets, the authors found that over time Mayne has positioned CA as an alternative to the counterfactual designs traditionally employed in evaluations, but that CA's epistemological basis and its stance on social causation also needs to be further elaborated.

During our assessment of CA in relation to the four typical weaknesses of theory-based evaluation approaches, the authors examined the use of CA in the context of two evaluations with very different research designs. The authors found that CA represents a useful approach for applying theory-based evaluation in practical evaluation work.

However, the CA methodology needs to be developed further in some respects in order to enable it to fulfill its potential. This was particularly found to be the case with regard to its approach to discerning and accounting for competing explanations and influencing factors. On the basis of our empirical application of CA, the authors proposed a number of revisions to the methodology.

ACKNOWLEDGEMENTS

REFERENCES

Bickman, L. (1987). The function of program theory: Using program theory in evaluation. *New Directions for Evaluation*, *33*, 5-18.

Bickman, L. (2000). Summing up program theory. *New Directions for Evaluation*, *87*, 103-112.

Chen, H.-T. (1990). Issues in constructing program theory. *New Directions for Evaluation, 47*, 7-18.

Chen, H.-T. (1994). Theory-driven evaluation: Needs, difficulties, and options. *American Journal of Evaluation*, *15*(1), 79-82.

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

Chen, H.-T., & Rossi, P. H. (1987). The theory-driven approach to validity. *Evaluation and Program Planning, 10*(1), 95-103.

Connell, J. P., & Kubisch, A. C. (1998). Applying a theory of change approach to the evaluation of comprehensive community initiatives: Progress, prospects, and problems. In K. Fulbright-Anderson, A. C. Kubisch, & J. P. Connell (Eds.), *New approaches to evaluating community initiatives, Vol. 2: Theory, measurement, and analysis* (pp. 15–44). Aspen, CO: Aspen Institute.

Cook, T. D., Scriven, M., Coryn, C. L. S., & Evergreen, S. D. H. (2010). Contemporary thinking about causation in evaluation: A dialogue with Tom Cook and Michael Scriven. *American Journal of Evaluation, 31*(1), 105-117.

Coryn, C. L. S., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A systematic review of theory driven evaluations practice from 1990 to 2009. *American Journal of Evaluation, 32*(2), 199–226.

Dahler-Larsen, P. (2001). From programme theory to constructivism: On tragic, magic and competing programmes. *Evaluation*, *7*(3), 331-349.

Davidson, E. J. (2000). Ascertaining causality in theory-based evaluation. *New Directions for Evaluation*, *87*, 17-26.

Donaldson, S. I. (2010). *Program theory driven evaluation science: Strategies and applications*. New York, NY: Lawrence Erlbaum Associates.

Fixsen, D., Naoom, S. L., Blasé, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: University of South Florida.

Hofmann, C.-A., Roberts, L., Shoham, J., & Harvey, P. (2004). *Measuring the impact of humanitarian aid: A review of current practice* (HP Report 17). London, UK: Overseas Development Institute.

House, E. (2001). Unfinished business: Causes and values. *American Journal of Evaluation, 22*(3), 309-315.

Hunter, D. (2006). Daniel and the rhinoceros. *Evaluation and Program Planning, 29*(2), 180-185.

Imas, L. M., & Rist, R. C. (2009). *Road to results: Conducting effective development evaluations*. Washington, DC: World Bank.

Kotvojs, F. (2006, September). *Contribution analysis: A new approach to evaluation in international development*. Paper presented at the international conference of the Australasian Evaluation Society, Darwin. Retrieved from http://www.aes.asn.au/conferences/2006/papers/022%20Fiona%20Kotvojs.pdf

Kotvojs, F., & Shrimpton, B. (2007). Contribution analysis: A new approach to evaluation in international development. *Evaluation Journal of Australasia, 7*(1), 27–35.

Lipski, U., & Ignagni, E. (2001). *Evaluation of nutritional interventions: A background document* (Report # 3). Toronto, ON: Prevention Unit, Division of Preventive Oncology, Cancer Care Ontario. Retrieved at https://ospace.scholarsportal.info/bitstream/1873/4888/1/10309032.pdf

Mackay, R., Horton, D., Dupleich, L., & Anderson, A. (2002). Evaluating organizational capacity development. *Canadian Journal of Program Evaluation, 17*(2), 121-150.

Mark, M. M. (1990). From program theory to test of program theory. *New Directions for Evaluation, 47*, 37-51.

Marquart, J. M. (1990). A pattern-matching approach to link program theory and evaluation data. *New Directions for Evaluation, 47*, 93-107.

Mayne, J. (1999). *Addressing attribution through contribution analysis: Using performance measures sensibly* (discussion paper). Ottawa, ON: Office of the Auditor General of Canada. Retrieved at http://dsp-psd.pwgsc.gc.ca/Collection/FA3-31-1999E.pdf

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* (ILAC Brief 16). Rome, Italy: Institutional Learning and Change (ILAC) Initiative. Retrieved at http://www.inteval-group.org/IMG/pdf/ILAC_Contribution_Analysis-2.pdf

Mayne, J. (2009). *Improving HRSDC evaluations of G&C program: Using theory-based approaches*. Unpublished manuscript.

Mayne, J. (2011). Addressing cause and effect in simple and complex settings through contribution analysis. In R. Schwartz, K. Forss, & M. Marra (Eds.), *Evaluating the complex: Attribution, contribution and beyond*. New York, NY: Transaction.

Newcomer, K. E., & Schreirer, M. A. (2001). *Using evaluation to support performance management: A guide for federal managers*. Arlington, VA: The Pricewaterhouse Coopers Endowment for The Business of Government. Retrieved from http://www.businessofgovernment.org/report/using-evaluation-support-performance-management-guide-federal-executives

Nielsen, S. B., & Ejler, N. (2008). Improving performance? Exploring the complementarities between evaluation and performance. *Evaluation*, *14*(2), 171-192.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Thousand Oaks, CA: Sage.

Ramboll Management Consulting. (2007). *Evaluering af forældrepålæ* [Evaluation of the parental injunction program]. Odense, Denmark: Danish Board of Social Services. Retrieved at http://www.servicestyrelsen.dk/om-os/filer/om-os/puljer-og-udbud/born-og-unge/57.pdf

Ramboll Management Consulting. (2010). *Evaluering af de utrolige år: Delrapport II* [Evaluation of the incredible years: Interim report II]. Odense, Denmark: Danish Board of Social Services. Retrieved at http://www.servicestyrelsen.dk/udgivelser/born-og-unge-alle-udgivelser/rapporter-og-boger-born-unge/evaluering-af-de-utrolige-ar-delrapport-2-effekten-af-programmet

Ramboll Management Consulting. (2011). *Evaluering af de utrolige år: Endelig rapport* [Evaluation of the incredible years: Final report]. Odense, Denmark: Danish Board of Social Services.

Rogers, P. (2007). Theory-based evaluation: Reflections ten years on. *New Directions for Evaluation*, *114*, 63-67.

Rogers, P., et al. (2000). Program theory evaluation: Practice, promise, and problems. *New Directions for Evaluation*, *87*, 5-13.

Schreirer, M. A. (1987). Program theory and implementation theory: Implications for evaluators. *New Directions for Program Evaluation*, *33*, 59-76.

Scottish Government. (2009). *Contribution analysis: Guide 6 in the Social Science Methods Briefings series*. Retrieved from http://www.scotland.gov.uk/Topics/Research/About/Social-Research/Methods-Guides/Guide-6

Smith, N. L. (1990). Using path analysis to develop and evaluate program theory and impact. *New Directions for Evaluation*, *47*, 53-57.

Weiss, C. H. (1997). How can theory-based evaluation make greater headway? *Evaluation Review*, *21*, 501-524.

White, H. (2010). A contribution to current debates in impact evaluation. *Evaluation, 16*(2), 153-164.

Winther, S. (2001). Implementeringsforskningen og dens relation til evaluering [Implementation research and its relationship to evaluation]. In P. Dahler-Larsen & H. Krogstrup (Eds.), *Tendenser i evaluering* [Trends in Evaluation] (pp. 61–78). Odense, Denmark: Odense Universitetsforlag.

Line Dybdal is manager at Rambøll Management Consulting in Denmark. She holds a M.A. research degree in social anthropology from the University of Aarhus, Denmark. During the last eight years she has been designing, planning, and conducting a number of evaluations within social affairs in Denmark. Her main research interest lies in theory-based approaches to evaluation.

Steffen Bohni Nielsen holds a Ph.D. from University of Aarhus, Denmark. He is currently Director of Evaluative Knowledge at Rambøll Management Consulting. For the past 10 years he has planned, conducted, and reviewed numerous evaluation studies within the human services in Denmark and abroad. He has also worked extensively with evaluation capacity building and results-based management. He is a former board member of the Danish Evaluation Society and a member of the International reseach group on policy and program evaluation (Inteval).

Sebastian T. Lemire is a senior consultant at Rambøll Management Consulting in Denmark. He holds an M.Sc. in social science methodology from the London School of Economics. He has carried out a number of evaluations, mainly in the educational sector in Denmark and the USA. His current research interest lies in methodologies that address the attribution question, including theory-based evaluation. He is a board member of the Danish Evaluation Society.