



Applying contribution analysis: Lessons from five years of practice

Evaluation
18(3) 281–293
© The Author(s) 2012
Reprints and permission: sagepub.
co.uk/journalsPermissions.nav
DOI: 10.1177/1356389012450810
evi.sagepub.com



Thomas Delahais

Eureval, France

Jacques Toulemonde

Eureval and Lyon University, France

Abstract

Contribution analysis (CA) has been advocated and discussed more than it has been applied. This article draws on five evaluations applying CA in the context of EU policies in the areas of development aid, agriculture, employment, and governance. Over the last five years, the authors have developed their capacity to implement the six steps of CA and especially to draft contribution stories and to make contribution claims. These practical efforts have, in turn, caused them to explore the foundations of the approach and refine and operationalize key concepts such as contribution claim, causal mechanism, and strength of evidence.

Keywords

causal claim, causal mechanism, contribution analysis, impact evaluation, logic model, theory-based evaluation, theory of change

Introduction

Contribution analysis (CA) is a pragmatic approach to applying the principles of theory-based evaluation. It assesses causal chains from beginning to end, reports on whether the intended changes occurred or not, and identifies the main contributions to such changes, including the intervention under evaluation. Over the last 10 years, CA has repeatedly been recommended in evaluation guidelines and has attracted interest at international events (Monnier, 2009; Toulemonde, 2010). However, instances of rigorous implementation have been surprisingly scarce and the theoretical foundations are still being strengthened (Mayne, this issue).

Corresponding author:

Thomas Delahais, Directeur adjoint Innovation & développement, 9, rue du château d'eau, F-75010 Paris, France.
Email: delahais@eureval.fr

This article draws on experience gained and questions raised in applying CA in a series of evaluations carried out by Eureval, a consultancy providing evaluation services to public authorities mainly in Europe. When we came across CA (Mayne, 1999, 2001), we were struck by: (1) how close it was to much of our current evaluation practice; and (2) its potential for strengthening the rigour and conclusiveness of theory-based evaluations (Rogers, 2008; White, 2009). Since 2005 (Toulemonde et al., 2011), we have used CA about 15 times in evaluation assignments in multiple areas such as agriculture, development aid, employment, environment, governance, and transport. In this article, we build upon five evaluations undertaken between 2007 and 2010, in which CA enabled us to assess impacts in a conclusive and useful manner while a counterfactual-based analysis would not have been feasible. However, we cannot yet qualify these evaluations as fully rigorous. The five examples of CA are presented below in successive boxes called *Policy Coordination*, *Development Aid*, *Employment*, *Active Citizenship* and *Agriculture*.

The first section describes some of the main lessons learned from the implementation of the six steps of CA. In the subsequent section, we explain how we operationalize some of the basic concepts used in the analysis.

Implementing the six steps of contribution analysis

This section discusses the six steps of CA (Mayne, 2008, 2011, and this issue). We had followed these steps more or less for years, as most of them are common to other forms of theory-based evaluations. Only Step 4 (drafting the contribution story) includes concepts that were new for us, and needed to be further operationalized. In this section following the six steps, we explain how the CA approach changed our previous practice, and highlight the hidden difficulties.

Circumscribing the causal issue

Our main approach to circumscribing the causal issue has always been to concentrate on a few cause-and-effect questions, and to specify these questions as precisely as possible in terms of scope (evaluated activities), and impact(s) to be assessed. There is frequent bargaining over causal issues between commissioners wanting to ask many broad questions and evaluators claiming that a small number of precise questions improves both the substantive quality and usefulness of the answers. In our five examples, the number of questions ranged from three to 14 per evaluation, of which two to five required a causality analysis. In two instances (*Policy Coordination* and *Development Aid*) we could cluster several questions into a single causal issue that was addressed through a single analytical approach. Finally, the number of causal issues in the five examples ranged from one to four.

Generally, we had no option other than CA for carrying out the impact analyses. Only in one instance (*Employment*) were usual attribution-analysis methods applicable (i.e. a counterfactual with matched comparison group and a synthesis of counterfactual-based evaluations). Initially, we were willing to apply CA to all remaining causal issues; for instance, four issues in the case of *Development Aid*. However, we gradually realized that this was not feasible. In fact, CA has proved for us to be more time-consuming than other theory-based approaches (Weiss, 1997). Moreover, our approach to applying CA requires significant stakeholder inputs at the inception phase (each causal issue requires an in-depth discussion) and at the fifth step of criticizing the draft contribution story (see below). Considering the current constraints applying to EC evaluations in terms of budget and stakeholder availability, we came to the conclusion that only one CA is manageable in a given evaluation assignment. We therefore tend to apply this approach to addressing questions that are particularly challenging, and answer other cause-and-effect questions with a conventional theory-based approach.

In one instance, the *Development Aid* evaluation (see Box 1), we realized how important circumscribing the causal issue was. In agreement with the reference group, we had decided to focus on Jordanian exports to the EU, although an alternative option was to consider Jordanian exports worldwide. This seemingly minor choice had deep consequences for the theory of change that had to include three causal chains in the first option and two chains otherwise. At the end of the evaluation, our assessment was negative (the intervention had made an insufficient contribution to the growth of Jordanian exports to the EU). Taking the alternative option would have led to a positive assessment (the intervention had made a satisfactory contribution to the growth of Jordanian exports to the rest of the world). This example shows that the causal issue should continue to be considered until an agreement is reached about its definition and delineation.

Box 1. *Development Aid* example.

This evaluation covered EU support to Jordan over a 10-year period (EC, 2007a). CA was applied to European support in the areas of trade and private-sector development.

The causal issue was not easy to specify since the EU had implemented various activities in this area with a number of objectives. The Evaluation Reference Group (ERG) decided that the success of the EU support should be defined as a breakthrough in Jordan's exports to the EU market. A logic model was developed in the form of three causal chains, connecting the main intended impact and respectively: (1) withdrawal of tariffs and quotas on Jordanian products, (2) technical assistance to Jordan's government and (3) free advisory services to exporting SMEs.

The available information was used to assess each link in the three causal chains, although with an uneven potential for cross-checking and triangulation. Other contributing factors were also explored.

The contribution story was drafted, discussed extensively in the ERG, and then strengthened. A first contribution claim is that Jordanian SMEs have improved their capability to export and that the EU support programme has made a major contribution to that change. However, a contribution of the same order was made by a similar US-funded programme. Another finding is that Jordanian exports to the EU have grown steadily but without any breakthrough.

This disappointing finding follows from a failure in the design of the evaluated policy: nothing was done to assist SMEs to match EU technical standards, a costly and difficult process which constitutes the main barrier to Jordanian products entering the European market.

Developing the theory of change

In line with theory-based evaluation, we have always dedicated significant resources to reconstructing the logic model of the evaluated intervention. We used to start by devising a diagram of intended impacts, but not a comprehensive theory of change including external factors and causal loops as recommended by Mayne (this issue). Our practice was to wait until the data-collection phase to explore, identify and assess the other contributing factors. In our five examples, the simplest model includes six changes symbolized by boxes and five causal links represented by arrows (*Employment*). The most complicated model includes 12 boxes and 16 arrows (*Policy Coordination*).

We have progressively come to understand that the efforts required to undertake a CA are proportionate to the number of boxes and arrows. If the logic model is too complicated (too many boxes, or causal chains are too intricate), the analysis may become unfeasible. In that case, it is probably a good idea to circumscribe a less complicated causal issue. On the other hand, the logic model should not be simplistic. Logical gaps should be bridged by adding intermediary boxes and arrows.

In the case of *Policy Coordination* (see Box 2) we developed a logic model comprising three causal chains in parallel. During the evaluation, we identified a rival explanation (EU coordination frames policy issues and then shapes reform agendas at national level). We also realized that this

rival explanation was not totally new for some stakeholders. To a certain extent, it was part of the implicit knowledge of the organizations concerned (Marra, 2004). As such, it could have been included in our model as a fourth causal chain. If we had unveiled this implicit organizational knowledge before moving to the next step of the CA, it could have been turned into formally validated knowledge, something that we could not achieve because the rival explanation was discovered too late in the analysis process. Hence, and in line with Mayne's views, we now consider that sufficient time and efforts should be invested in documentary reviews and exploratory stakeholder interviews at the second step of the analysis.

Box 2. *Policy Coordination example.*

The evaluation, carried out in 2007, assessed the EU instrument used for coordinating national policies in the areas of economy, labour market, and environment (EC, 2007b). We applied CA for addressing the question of whether the coordination instrument had helped to introduce reforms on the political agendas in EU Member States.

We first built a logic model eliciting the changes and causal links as they were understood by stakeholders at the time of the evaluation. We then gathered secondary data plus additional information obtained through: (1) interviews with policy-makers, experts and stakeholders in 18 countries, and (2) case studies covering four policy areas and 12 countries.

Our evaluation confirmed that the coordination instrument contributed to fostering mutual learning and shaping reform agendas, but in conjunction with a number of other factors that were making stronger contributions. The evaluation also shows that the 'peer pressure' mechanism, which was part of the logic model, was not working. Moreover, doubts were cast on the effectiveness of the pressure of public opinion, another causal chain in the logic model.

The CA revealed a mechanism called the 'framing of policy issues', which is not part of the initial logic model. This mechanism was found to make an important contribution to national reform agendas, although in conjunction with other major contributors such as multilateral institutions and knowledge communities.

Our practice of developing logic models also exacerbated a problem that is probably common to all theory-based evaluations, i.e. an unbalanced attention paid to the causal links under test at the expense of other contributing factors and rival explanations. Such a bias tends to affect the understanding of the causal mechanisms, the gathering of evidence, and interpretations. It is not just a positive bias or a confirmation bias (Oswald and Grosjean, 2004) since we may be driven to refute the model as well as to confirm it. The risk is rather to have our evaluation shaped by and bounded in a narrow logic, and unable to identify alternative explanations and to account for these. In order to prevent that risk, we came to recognize the need for a systematic search for potential alternative explanations before gathering evidence, rather than during the data collection stage. In this sense, we are now on the same wavelength as Mayne who recommends (this issue) that a comprehensive theory of change be developed at the second step of CA (see also Lemire et al., this issue).

Gathering evidence

According to Mayne (this issue), the third step of CA is devoted to gathering available evidence from secondary data and experts' views. Additional evidence from primary data is to be collected later on at the fifth step (strengthening the contribution story). On the contrary, our practice has always been to collect primary and secondary data in parallel, something that better matches the

tight time constraints of our assignments. We used to build up a data-collection work plan and to apply a range of tools enabling us to answer each evaluation question on the basis of several distinct information sources, in line with the principle of triangulation.

Our experience with CA entailed a major change in the sense that we now seek to apply the triangulation principle to each step of reasoning and not just to each evaluation question. For instance, in the *Development Aid* example, the theory of change was composed of 10 boxes, nine arrows, and about 20 other contributing factors. Some secondary sources were available from previous evaluations and studies, but they enabled us to cover less than 50 per cent of the theory of change, with a very limited prospect for triangulation. Our data collection work plan included interviews with policy-makers and experts, and a series of visits to assisted enterprises. Thanks to these tools, the evidence gathered covered 90 per cent of the theory of change, but only a small minority of our reasoning arguments were supported by triangulated sources. This case exemplifies the fact that our approach to CA is quite demanding in terms of information sources.

Progressively, we have found that embedded case studies are excellent at providing evidence that covers the whole theory of change. Embedded case studies start as usual case studies and supply rich information about the first logical links of the theory of change. Then sub-cases are selected within each case and investigated in order to inform on further logical links, and so on until the whole theory of change is covered. An example of embedded case study design is displayed in Box 3 (*Employment*). In this example, our evidence base originated from four distinct countries, which offered a good prospect for triangulation.

Box 3. Employment example.

A study used for reforming the European Social Fund (EC, 2010), asked how much policy makers learn from rigorous impact analyses based on counterfactuals. We addressed this question through a meta-evaluation using a CA.

First, we developed a logic model that stemmed from our own knowledge of how evidence percolates in the policy-making sphere and we then investigated four success stories of impact evaluations carried out between 2005 and 2009. It was difficult to gather evidence about the political use (or non-use) of evaluative knowledge, but we managed to generate the information we needed through four embedded case studies. The four cases were high-quality impact evaluations. In each case, we selected a specific finding of interest to policy-makers (sub-case), and then a circumstance in which that finding was used or could have been used for policy-making purposes (sub-sub-case).

The four exemplary evaluations delivered new, credible, and relevant lessons, but learning from these lessons was limited because: (1) evaluation findings were pushed by information suppliers much more than they were pulled by policy makers, and (2) technical issues were given too much attention at the expense of substantive findings. Knowledge tended to be used indirectly, feeding into streams of successive evaluations and studies until they were converted into problems, explanations, and solutions that were useable in the policy-making process.

Even with a relatively simple theory of change including just one causal chain, as in the above example, the reasoning arguments are numerous and the amount of information needed to support all arguments is considerable. Initially, we thought that CA might be a less expensive alternative to counterfactual-based methods, but we have now understood that our approach is also demanding in terms of data collection, at least if we claim to be rigorous. We now consider that any state-of-art impact analysis is costly, irrespective of whether it relies on an attribution or a contribution approach. This is why we do not feel capable of addressing more than one causal issue rigorously in a typical EU evaluation.

In the case of *Development Aid*, we gathered a wide range of data, from which we extracted 70 items of evidence (see five of these in Table 1). Two items out of three were confirming the logic model. This proportion may just mean that the intended changes tended to occur, that the causal mechanisms tended to work, and that other contributing factors were marginal. We might also suspect that investigators were unconsciously biased toward confirming the logic model because we had not developed a comprehensive enough theory of change (see above). A positive bias might just result from the fact that data collection tends to be entrusted to the most junior members of the evaluation team or to sector experts who may not be that familiar with evaluation.

We have tested several approaches in order to reduce such risks:

- In the case of *Agriculture* (see Box 4), the draft case monographs were reviewed by an external expert, to make sure that whenever possible, the authors had sought information confirming and refuting the logic model;
- In the case of *Active Citizenship* (see Box 5), we convened several telephone conferences during the case-study investigations, in which the views of the investigators were systematically challenged in order to avoid positive biases.

Box 4. Agriculture example.

The evaluation assessed the EC financial support to the promotion of high-quality European agricultural products (e.g. wine, cheese, vegetables, etc.) towards India and China (EC, 2008), with special attention paid to the impact on food professionals and consumers. We first built a logic model mainly based on our knowledge of similar schemes, and then tested this model through a series of 12 case studies. An independent expert assessed each case monograph in order to ensure that all causal assumptions were covered and that contextual factors were accounted for.

The main finding was that promotion activities targeted at professionals (rather than consumers) proved the most effective by creating sustainable professional relationships, which themselves contributed to the development of distribution networks and, in some cases, triggering limited (yet observable) increases in sales.

Box 5. Active Citizenship example.

From 2007, the EC supported two programmes aimed at encouraging citizens to debate European issues, through ‘citizen consultations’ run by selected NGOs (EC, 2009). The debates were intended to make citizens’ voices heard in EU policy-making processes. In evaluating these programmes, we used CA to answer a question about ‘the contribution of the programme to citizens’ debates on the future of the EU and the impact of the EU on their daily lives’.

We developed a logic model based on a literature review and expert views on what could reasonably be expected from the programme. The model was tested through 21 case studies that attempted to balance evidence confirming and disconfirming each intended contribution in the logic model, and to explore all other contributory factors. As the draft contribution story was quite negative, it was strongly challenged by the programme managers. This meant that several findings had to be either consolidated by further evidence or reformulated. Finally, despite enthusiasm among most of the participants about the deliberative process itself, programmes failed to trigger any debates outside their own small audience, contrary to the expectations in the promoters of the programme. They failed to obtain mass media coverage, and therefore had no influence on public opinion. Moreover, the outputs of the projects have not been used in policy-making because the politicians who attended the events were involved only superficially, and no clear status was given to the outputs of the debates.

Table 1. A record in the Evidence Analysis Database (example).

Label	Supported managers report a high increase in exports
Statement	In a questionnaire survey, managers report that the total annual sales changed by an average of 32% in the period covered by the European support. Exports increased by 61%.
Source	Survey contracted out by programme managers, and answered by almost all supported SMEs.
Type of source	Primary / Secondary
Causal link	Arrow 6–9 (SMEs that have been successfully assisted improve their competitiveness)
Confirming/refuting the logic model	Confirming / Refuting
Type of causal mechanism	Intended contribution / Other contribution / Condition to intended contribution / Intended condition to other contribution / Feedback*
Strength of evidence	Strong - Rather strong – Rather Weak – Weak*

*these two lines are further explained in the next section.

Overall, the process of gathering information for CA relies on the same principles and faces the same risks as with any other theory-based approach. What changes is the fact that CA is particularly demanding since the evidence base needs to cover all causal links, each one including a box, an arrow, and several contributing factors. If the analysis traces all reasoned arguments in a systematic way, then the gaps in the evidence base are easily visible.

Drafting the contribution story

This is the core step where CA adds most value. Successive attempts at applying CA led us to recommend an approach where data are aggregated no longer at the level of evaluation questions or criteria, but at the level of each causal link in the logic model. To facilitate such a systematic approach, we compile an Evidence Analysis Database which records and qualifies all items of evidence extracted from the set of information that has been gathered (see an example in Table 1).

All items of evidence related to a given causal link are clustered in an Evidence Analysis Table as exemplified in Table 2.

Table 2 helps to draft a contribution claim composed of a series of change statements and causal claims. For a given causal link, we identify up to three other contributing factors or conditions that are assumed to have a particular influence. We then review and interpret the related items of evidence and we write a contribution claim (see Table 3), i.e. a text starting by a change statement, describing the main causal mechanisms at play and ranking their influence by order of magnitude from first to last. The robustness of a contribution claim depends on the supporting items of evidence that may or may not be strong, convergent, and triangulated. The claim is said to confirm the logic model if the intended change occurred and if the intended contribution is highly ranked in comparison with other contributing factors.

All contribution claims are ordered along the boxes and arrows of the logic model, and assembled into the contribution story. In this typically 300-word text, each claim is connected to a detailed explanation, and then to the corresponding Evidence Analysis Table through hyperlinks.

Table 2. Evidence Analysis Table example.

Arrow 6–9 (assisted SMEs improve their competitiveness)				
Items of evidence	Type of source	Confirming/refuting	Causal mechanism*	Strength of evidence
Supported managers report a high increase in exports	Secondary	Confirming	Intended contribution	Rather strong
Clear examples of increased exports in supported SMEs	Primary	Confirming	Intended contribution	Rather weak
Interviewees cast doubts on the sustainability of benefits in a context of reinforcing competition	Primary	Refuting	Condition to intended contribution	Rather weak
USAID support programme with same target, magnitude and reported success	Secondary	Refuting	Other contribution	Rather strong
The opinion of experts and visited SMEs is that the business environment has not improved significantly	Primary	Confirming	Other contribution	Rather weak

*see explanation below.

Table 3. Contribution claim example.

Arrows 6–9 (SMEs that have been successfully assisted improve their competitiveness)

The competitiveness of supported SMEs on international markets has clearly improved, although in a context of reinforced competition that may soon put that success at risk. The EU-funded assistance and a similar USAID programme have made the first largest contributions to that change. No other major contribution has been identified, including in terms of business environment in Jordan.

Up to now all the elements of this process have been applied in different evaluations but not in a single one. This is mainly because particular innovative methods arose in different evaluation assignments. However, there is no specific difficulty in applying the overall approach, provided that: (1) sufficient resources are devoted to CA, and (2) users are receptive to evaluation reports written in this way.

Strengthening the draft contribution story

Our usual practice at this stage of the evaluation process was to submit our draft findings and arguments to:

- a careful review by the evaluation team leader;
- a quality assessment by a senior partner/expert who had not been part of the evaluation; and
- discussions in one or several meetings with representatives of the commissioning body, members of the Evaluation Committee, participants in a stakeholder workshop, or experts on a panel.

At the fifth step of a CA, the draft contribution story should be challenged through the same kind of open and critical discussion, preferably involving people who are distant enough from the evaluation team, whose values are sufficiently diverse, and who hold opposing stakes in the evaluation. Such a 360° scrutiny would be the ideal way to identify the weaknesses in the draft contribution story (Patton, this issue). However, we discovered that the process is constrained by the availability of stakeholders to review and discuss our analysis at a time when they have to absorb many findings, conclusions, and often recommendations. In the example of *Development Aid*, our 95-page draft report ended in 10 conclusions and nine recommendations. When we came to discuss this draft report in a two-hour meeting of the Evaluation Committee, most of those attending had read only that section of the report in which they had a stake. The specific conclusions derived from our CA were given a few minutes in a discussion that involved a few participants only. This was clearly not enough for reviewing the contribution claims in sufficient depth.

Moreover, some participants in the evaluation meetings do not feel competent to criticize the contribution story in detail. This happened to a large extent in the final meeting of the Evaluation Committee of the *Active Citizenship CA*. In three of our five examples, however, we received challenging and useful inputs from respectively institutional partners, evaluation stakeholders and the evaluation manager in the commissioning body.

Stakeholders who criticize the draft contribution story also provide additional information sources that may lead to additional data. For that purpose, resources were set aside that could be used for strengthening the analyses at this stage. Such additional evidence can easily be inserted into the Evidence Analysis Database and Tables, and quickly used for amending our interpretations, reasoning arguments, and contribution story.

More problematic is the occurrence of having to modify the theory of change at this step of the analysis. For instance, in the case of *Policy Coordination*, we found that we had to consider a new causal chain ('framing of policy issues'). Similarly, in the case of *Development Aid*, an unforeseen condition appeared to play a key role ('technical barriers to Jordanian exports at the entry of the EU market'). In both instances, the time and resources left did not allow us to develop and implement the kind of comprehensive data-gathering work plan that would have been required to undertake a systematic test of the new assumptions.

Finalizing the contribution story

Our usual approach to finalizing an evaluation was to present each conclusion, and the main arguments supporting it, in a section of the main report, which is itself reflected in a paragraph of the executive summary.

CA, however, is more demanding in terms of structured presentation, as its validity relies on a systematic coverage of all boxes and arrows along the causal chains, something that creates a risk of losing an overview of the big picture. Time and resources permitting, our solution to this problem has been to deliver an evaluation report in two volumes. The first one includes the executive summary and a user-oriented presentation of the findings, conclusions, lessons learned, identified problems, and suggested solutions. A second volume displays the full chain of reasoning and an appendix includes the evidence analysis database. The executive summary includes the contribution story or several such stories if more than one causal issue is addressed through a CA.

Contribution stories are written as short texts (300–500 words), which are easily accessible to public managers. However, such texts are too long for being used in public debates. In comparison, the findings of a counterfactual-based analysis are typically expressed in one sentence including a quantitative impact estimate that may be easily understood by both policy makers and journalists.

How can a contribution story be sold in the political arena? From our point of view, this is a problem that remains to be solved.

Operationalizing key concepts

As noted in the previous section, we consider that CA adds most value to our practices, and more generally to evaluation methods, when it comes to drafting the contribution story. However, we have found it particularly difficult to operationalize this fourth step of the analysis. The challenge is to make contribution claims that are based on evidence in a way that is rigorous, traceable, and credible. Even if this challenge still lies ahead, we have gained a much deeper understanding of the key concepts that frame our reasoning about contribution claims, i.e. causal mechanism, causal claim, and contribution claim. These concepts are illustrated below with reference to the *Development Aid* case.

Causal mechanism

Some authors define a mechanism as a constellation of causes that regularly bring about a particular effect (Astbury and Leeuw, 2010; Gerring, 2010; Hedstrom, 2005). In 1997, Pawson and Tilley considered that programmes generate their effects through various change mechanisms interacting with contextual factors. In line with these authors, we define a causal mechanism as a substantial explanation of why and how a given factor contributes to a given change. Each arrow in the logic model is a causal mechanism that explains an intended contribution to an intended change. It is itself part of a causal configuration including other mechanisms that explain, for instance, the influence of other contributing factors. We have encountered the following types of causal mechanisms:

- Intended contribution, i.e. arrow of the logic model, assumed to be sufficient for generating the intended change;
- Other contribution, i.e. influence of a contextual factor, also assumed to be sufficient for generating the intended change;
- Condition to intended contribution, i.e. influence of a contextual factor that is necessary for the intended contribution to work; Table 2 shows an example of such a condition: improved competitiveness is conditional on the fact that competitors make slower progress than supported SMEs;
- Intended condition to other contribution, i.e. a particular case in which the evaluated intervention is designed to block or unblock the influence of a given contextual factor;
- Feedback, i.e. reverse contribution.

All causal mechanisms associated with a given link in the logic model constitute a causal package (Mayne, this issue). A contribution is unnecessary but sufficient for the occurrence of the intended change. A condition is necessary for the working of a mechanism that is itself unnecessary but sufficient for the occurrence of the intended change (Mackie, 1974).

Causal claim

A causal claim asserts that a given change (or no change) is one of the causes of another change (or no change). For inferring such a claim from evidence, the analyst has to consider other influencing factors in order to eliminate them (Scriven, 2005) or to account for their contribution.

A causal claim may apply to all or part of the theory of change (e.g. the EU support has not boosted Jordanian exports to the EU market). In this section, we focus on causal claims that apply to one of the mechanisms associated with a given link in the logic model (e.g. newly acquired management skills have increased the competitiveness of supported SMEs). In the process of drafting a contribution claim, we proceed by compiling causal claims that apply to all mechanisms of a causal package.

CA welcomes causal claims inferred through all types of methods such as a probabilistic approach based on a statistically representative sample (e.g. Table 1), a comparative approach based on a number of case studies (e.g. Box 2), a triangulation approach based on the cross-checking of three independent information sources, or even a single item of evidence if it is assessed to strongly validate a reasoning argument (Greene, 2011).

Contribution claim

We understand that a contribution claim asserts that an intended change: (1) did or did not occur, (2) due or not due to the intended contribution, (3) in conjunction with a few selected contextual factors, (4) all considered mechanisms being explained and ranked by order of influence, and (5) other non-selected mechanisms being acknowledged.

A contribution claim may account for a simple causal package comprising just the intended contribution, a few other selected contributions, and the mention of some other eliminated mechanisms (e.g. Table 3). However, contribution claims may become difficult to write and to read if they include feedback loops and/or conditions (Sridharan and Nakaima, this issue).

Contribution claims often include statements about the magnitude of the causal relationship, expressed with terms such as ‘major’, ‘minor’, ‘marginal’, ‘important’, and so on. Building upon our practice, we consider that it is often possible to sort out the most influential factors by order of relative contribution, and to say that a given cause made the ‘third main contribution’, or an ‘even contribution as’. However, our practice in this respect is still empirical and not fully conceptualized.

Conclusions

Theory-based approaches are good at explaining impact mechanisms, but often inconclusive as to whether interventions do or do not work. CA has the potential for both concluding on the achievement of impacts and providing evaluation users with the explanations they need. Moreover, it has the potential to do this where no counterfactual-based method is possible, as shown by the five examples illustrating this article.

Theory-based evaluations are often challenged on the grounds that they do not produce rigorous findings as far as causes and effects are concerned. CA has the potential for changing such a vision because its six steps can be explained and implemented in a structured and traceable manner. However, our approach to applying CA is demanding in terms of resources and competencies, something that means that we cannot promise to deliver a rigorous analysis of more than a few causal issues in a typical EU evaluation.

In order for a CA to be acknowledged as rigorous, we would need quality criteria and benchmarks to be developed, applied, and recognized in the evaluation profession. Setting strict quality standards is a precondition for CA to compare with other impact evaluation methods. We may imagine that CA practitioners could reach such consensus in the future, but the way ahead is obviously long. All six steps of the CA should be associated with specific quality criteria, and

especially Step 4 (drafting the contribution story) where CA adds most value to the current practices and where the analysis is most likely to be challenged.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

- Astbury B and Leeuw F (2010) Unpacking black boxes: mechanisms and theory building in evaluation. *American Journal of Evaluation* 31(3): 363–81.
- EC (2007a) Evaluation of the European Commission's support to the Hashemite Kingdom of Jordan. URL: http://ec.europa.eu/europeaid/how/evaluation/evaluation_reports/2007/1087_docs_en.htm.
- EC (2007b) Evaluation of the Integrated Guideline Package (IGP) for Growth and Jobs. URL: http://ec.europa.eu/economy_finance/evaluation/pdf/final_report_ipg_en.pdf.
- EC (2008) Evaluation of promotion measures of European agricultural products in third countries (not available online).
- EC (2009) Evaluation of the Plan D / Debate Europe citizen consultation projects. URL: http://ec.europa.eu/dgs/communication/pdf/evaluation_citizens_consultations_sep2009_en.pdf.
- EC (2010) Study on the Return on ESF Investment in Human Capital. URL: <http://ec.europa.eu/esf/BlobServer?docId=160&langId=en>.
- Gerring J (2010) Causal mechanisms: yes, but ... *Comparative Political Studies* 43(11): 1499–526.
- Greene J (2011) The construct(ion) of validity as argument. In: Chen HT, Donaldson SI and Mark MM (eds) *Advancing Validity in Outcome Evaluation: Theory and Practice. New Directions for Evaluation* 130. San Francisco, CA: Wiley, 81–91.
- Hedstrom P (2005) *Dissecting the social: on the principles of analytical sociology*. Cambridge: Cambridge University Press.
- Lemire ST, Nielsen SB 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 (1974) *The Cement of the Universe: A Study of Causation*. Oxford: Oxford University Press.
- Marra M (2004) The contribution of evaluation to socialization and externalization of tacit knowledge: the case of the World Bank. *Evaluation* 10(3): 263–83.
- Mayne J (1999) Addressing attribution through contribution analysis: using performance measures sensibly. Discussion paper. Office of the Auditor General of Canada. URL: <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 No. 16: Rome: The Institutional Learning and Change Initiative. URL: http://www.cgiar-ilac.org/files/publications/briefs/ILAC_Brief16_Contribution_Analysis.pdf.
- Mayne J (2011) Contribution analysis: addressing cause and effect. In: Schwartz R, Forss K and Marra M (eds) *Evaluating the Complex*. New Brunswick, NJ: Transaction Publishers, 53–96.
- Mayne J (2012) Contribution analysis: Coming of age? *Evaluation* 18(3): 270–280.
- Monnier E (2009) Designing a contribution analysis for assessing expected impacts. Workshop material for the Cairo Conference on *Perspectives on Impact Evaluation*. URL: <http://www.3ieimpact.org/userfiles/file/W-15.zip>.
- Oswald ME and Grosjean S (2004) Confirmation bias. In: Pohl RF (ed.) *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory*. Hove: Psychology Press, 79–96.
- Patton MQ (2012) A utilization-focused approach to contribution analysis. *Evaluation* 18(3): 364–377.
- Pawson R (2002) Evidence-based policy: the promise of 'realist synthesis'. *Evaluation* 8(3): 340–58.
- Pawson R and Tilley N (1997) *Realistic Evaluation*. Thousand Oaks, CA: SAGE.

- Rogers PJ (2008) Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation* 14: 29–48.
- Scriven M (2005) Can we infer causation from cross-sectional data? Paper for the National Academy of Sciences. URL: http://www7.nationalacademies.org/bota/School-Level%20Data_Michael%20Scriven-Paper.pdf.
- 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): 378–395.
- Toulemonde J (2010) Evaluating impact through a contribution analysis. Workshop at the Prague Conference of the *European Evaluation Society*. URL: http://www.ees2010prague.org/images/file/W404_Evaluating%20Impact%20through%20a.pdf.
- Toulemonde J, Carpenter D and Raffier L (2011) Coping with the evaluability barrier: poverty impact of European support at country level. In: Forss K, Marra M and Schwartz R (eds) *Evaluating the Complex: Attribution, Contribution, and Beyond*. New Brunswick, NJ: Transaction Publishers.
- Weiss CH (1997) How can theory-based evaluation make greater headway? *Evaluation Review* 21: 501–24.
- White H (2009) Theory-based impact evaluation: principles and practice, international initiative on impact evaluation. URL: <http://www.3ieimpact.org>.

Thomas Delahais, is a consultant in evaluation at Eureval. He teaches at the Master Programme in Monitoring and Evaluation of the University of Lyon.

Jacques Toulemonde, evaluation expert, is partner and co-founder of Eureval and associate professor at the University of Lyon, where he manages a Master Programme in Monitoring and Evaluation.