

Achieving impact: impact evaluations and narrative simplification

Tomas Hellström^a and Christina Hellström^b

^aSchool of Economics and Management, Lund University, Lund, Sweden; ^bDepartment of Psychology, Göteborg University, Göteborg, Sweden

ABSTRACT

This study is concerned with how impact from research and innovation (R&I) programmes is accounted for in impact evaluation reports. Establishing causal links between a research funding instrument and different effects, poses well known methodological difficulties. In the light of such challenges, textual accounts about causal links ought to be carefully written. Nevertheless, impact evaluation reports have a tendency towards unwarranted simplification as far as impact inferences are concerned. In this study, we illustrate how such simplifications – versions of the narrative device ellipsis – are accomplished. Using examples from three Swedish impact evaluation reports, we focus on the constituent components of longer impact accounts, that of the impact argument, to analyze the various ways that impact is narratively achieved through simplification. We believe this analysis can contribute to the methodology of impact evaluation, as well as spread light on some the difficulties in the historiography of innovation in general.

Introduction

There is an increasing, albeit longstanding, interest in trying to account for the social and economic impacts of research, and specifically the effects of research funding for a variety of reasons. Accountability demands and ‘value for money’ pressures have been on the rise for a while, but the trend can also be associated with more comprehensive research and innovation (R&I) policies in general. Such reasons as providing an overview of a national research system, informing funding decisions, and enabling knowledge utilization and technology transfer belong to the latter category (Penfield *et al.*, 2014). The ‘impact movement’ is today most notably reflected in UK’s Research Excellence Framework (REF), where institutions describe their approach to impact and provide case studies exemplifying previous impacts (REF, 2014). The ‘Payback Framework,’ developed to account for benefits from health and biomedical science (Wooding *et al.*, 2005; Donovan and Hanney, 2011) is another well known example, and both have had many spin-offs and adaptations (e.g., Donovan, 2011).

This article focuses on one expression of impact assessment: the programme impact evaluation, and especially how such evaluations are presented in support of specific

programmes. Programme impact evaluations, using a mix of qualitative and quantitative approaches are undertaken to establish to what extent the funding from an agency programme has resulted in socio-economic impacts and other outcomes (National Academy of Sciences, 1999). In the case of research programmes, such programme evaluations are typically conducted a few years after programme termination to allow for impacts from the research to manifest themselves, such as the adaptation of results into new technologies, products, processes or other practices. While there has been some recent work done on research impact at the level of individual researchers (Gläser and Laudel, 2015), the issue of impact at the agency level has a much longer history. Establishing the link between funding and impact has been an imperative for funding agencies, especially mission agencies whose sole purpose is to further improvements in some sector. However, inferring a connection between research funding and impact is tricky. Kostoff (1993), in reviewing the classical retrospective impact studies Project Hindsight (sponsored by the US Department of Defense in 1965) and Project TRACES (by the National Science Foundation in 1968), noted several ways in which the path from research to innovation resists analysis. First, there is the issue of time. Innovation from basic research frequently occurs several decades after a discovery or publication. In the case of applied, directed research the time is closer to one decade (see also Mansfield, 1998). From a methodological perspective a more important issue is the problem of accounting for all those long-term cumulative factors that bear on innovations. Rather than emanating from specific identifiable results in research, innovation typically seems to draw from several complementary sources of research and technological development; in fact, all innovations in Project Hindsight and TRACES depended for their realization on such complex 'pools of knowledge' (Kostoff, 1993).

The methodological challenges of capturing the connection between intervention and impact are easy to acknowledge. However, while the likelihood of accounting for the genuine research precursors of any significant innovation in simple causative terms is probably slim, such accounts are routinely produced and used in science policy at an accelerated speed. To address this issue, we focus on what might be called the 'discursive challenges' of producing convincing accounts of these connections. Having positioned the article's contribution in this way, we do recognize that the two challenges are closely related. Methodological difficulties, for example, are more than likely to appear as presentational ones in a final evaluation report. The critical question here, and the one we pursue, is whether discursive inferences are supported on the level of argument. Are we given compelling reasons to believe in them? This article takes its point of departure in the apparent need to simplify the historiography of innovation in order to validate (justify) previous funding from a socio-economic (political) perspective. Specifically, we are interested in the way such evaluations are narrated to 'achieve' a sense that impact from some specific agency funding has actually taken place. Not unexpectedly, this is done by means of simplification, where important parts of the descriptive historical account are omitted, a narrative device referred to as 'ellipsis' in literary theory, and which is elaborated in this article.

By explicating and exemplifying the narrative devices used to 'achieve impact' in this way, we hope to promote a more critical stance towards these accounts, and also to point towards some inherently tricky aspects of the historiography of innovation. In

doing this, our study draws on a number of impact evaluations published by a Swedish agency responsible for promoting innovation on a national level. These evaluations cover several fields of science and innovation, and are specifically intended to account for the impact from three programmes of applied research funded by the agency and its predecessor. The article is structured as follows: first, we review some methodological challenges involved in impact evaluations generally, as well as discuss the use of textual, ‘argumentative analysis’ in policy research, with special reference to narrative analysis. Second, we elaborate on the material and method employed in this study, following this with a section presenting examples of impact accounts displaying various forms of ellipsis. The article ends with some conclusions on the significance of narrative devices in research policy texts, and how one might improve impact evaluations in the light of our analysis.

Impact evaluations

Impact evaluation, or impact analysis, is the attempt to map the effects of policies and programmes. It is a specific form of outcome evaluation, often applied on a programme level. Mohr (1995, p.1) defines it as ‘determining the extent to which one set of directed human activities (X) affected the state of some objects or phenomena ($Y_1 \dots Y_k$) and ... determining why the effects were as small or large as they turned out to be.’ This account is a purely descriptive one, not taking any other aspects of the programme into consideration than the way that its activities furthered the programme goals. This is the most common way of thinking about impact evaluation (Smith and Larimer, 2017). A more normative take is that of Weiss (1998), who suggests that impact evaluation should also consider the programme/policy in relation to a number of explicit or implicit standards; for example, the worth of the programme in advancing desirable social goals, and how policy can be improved in the light of the results.

Impact evaluations should as a minimum connect a programme’s activities to some effects and explicate how these effects relate to the goals of the programme, or the problem that the programme is intended to solve (White, 2010). Sometimes the outcome that the programme addresses is referred to as its ‘outcome of interest.’ For example, in attempting to raise international competitiveness in technology-based small and medium-sized enterprises (SMEs) by improving their product development processes, the first is likely to be the outcome of interest, and the second an intermediary outcome. The outcome of interest has to: (1) operationalize a key aspect of the problem; and (2) be amenable to being causally linked to the programme (Smith and Larimer, 2017). The first of these is mainly a normative, conceptual issue, and the second is methodological, involving challenges of causal attribution. Both the normative and the methodological operationalization of the outcome of interest are problematic. The normative because policies and programme goals are commonly unclear and ambiguous; they are often decomposable into several problems and some of these may turn out to be contradictory in terms of the actions needed to address them (Hogwood and Gunn, 1984).

The second methodological challenge usually involves various aspects of attributing causal connections between activities and effects in the absence of proper experimental conditions (i.e., randomly-assigned treatment and control group designs). The challenge

of determining the effect by establishing what would have happened without the intervention is known as the counterfactual problem in impact evaluation (Gertler *et al.*, 2016). Since most policy programmes do not allow a random assignment of the ‘treatment’ to the target population or an equivalent control group, a robust estimation of the counterfactual is often impossible. Instead, the evaluator comes to rely on other means, such as tracing the processes presumably leading from activities to effects (*cf.* ‘process tracing’), often utilizing some kind of logic model depicting (*ex ante*) how the programme yields various outputs and effects (O’Keefe and Head, 2011). Such logic models or programme theories rely on the validity of the causal attributions between their component parts (e.g., inputs → activities → outputs → outcomes → impacts). While much has been made of the ‘evidence base’ for establishing generalizable connections, the track record of this evidence base, especially the transferability of the results of ‘randomized controlled trials’ (RCTs) into untried contexts, is questionable on empirical as well as conceptual grounds (e.g., Cartwright and Hardie, 2012). These issues have generated heated discussion in evaluation research regarding how to warrant causal claims, or whether to make them at all (see Guba and Lincoln, 1989 for an early statement of the second position). Scriven (2008) is one among several commentators who reject the notion that RCTs confer knowledge of causal relations in the above counterfactual sense. Arguing that RCTs are not strictly experimental, he defends a qualitative, single case approach to evaluation that takes contextual rather than generalizable mechanisms into account (see also Stern, 2013). Today, the discussion on impact evaluation centers on the tension between the generalizability of the relationships mediating between interventions and effects (e.g., Cartwright, 2007; Scriven, 2008), and the role of theory in formulating explanations for programme impacts that arise from ‘realist evaluation’ (Pawson and Tilly, 1997).

The result of this discussion is a wider available spectrum of positions on how to view causal claims and what to expect from causal arguments (Gates and Dyson, 2017). For present purposes, it is important to note that the absence of universally accepted ways of translating social goals into operational programme goals (the normative challenge), and stringent or even minimally acceptable measures for capturing the counterfactual (the methodological challenge) force the analyst into a series of informal or intuitive accounts. These involve assuming the social and political validity of certain operational programme outcomes/effects, and (spuriously) selecting indicators to represent such goals, as well as basing inferences on ad hoc arguments for how activities can lead to effects. This is usually done by applying logic models and programme theory, which themselves are often informal arguments in the form of chains of propositions resting on black-boxed or untested assumptions. The use of discursive ploys, such as narrative devices that allow for simplification of these arguments, thus obscuring vital facts and inferences, easily finds its way into such analyses. For this reason, impact evaluations are eminently amenable to what has been referred to as argumentative analysis in policy studies (Fischer and Forrester, 1993).

Policy and argumentation

The argumentative turn in policy analysis puts the emphasis on how policy texts accomplish their effects through persuasion, rhetoric and various narrative devices (e.g., Fischer and Gottweis, 2012). A central tenet of this tradition is reflected in

Majone's (1989) suggestion that policy analysts 'seldom demonstrate the correctness of their conclusions, but only produce more or less persuasive arguments' (p.42). The policy analyst becomes a producer of arguments rather than facts; these arguments rely, more or less tacitly, on norms and contestable ultimate goals, and their purpose is to produce policy action rather than intellectually agreed insight. According to Majone, policy analysis, via reports and advice, provides standards for public debate. It can therefore be evaluated normatively from the point of view of how those texts conform to standards for public reason.

One set of standards center around to what extent we are prepared to accept a rhetorical component in policy analysis; for example, an amount of rhetorical simplification, or the withholding of key elements in an argumentative sequence that might not hold up to closer scrutiny (Majone, 1989; Roe, 1994). One such strategy is enacted in the attribution of a causal relationship (cause and effect sequences) on an historical fragment, by the way in which narrative order represents events. Seymour Chatman suggests that such narrative sequencing can be so persuasive that we do not even need to have causes and effects explicitly connected to accept them as such (Chatman, 1978). Chatman writes that 'The interesting thing is that our minds inadvertently seek structure, and will provide it if necessary ... we do so in the same way we seek coherence in the visual field, that is, we are inherently disposed to turn raw sensation into perception' (pp.45–46). It may be exactly this quality of narrative, its implicit connection of events into a causal sequence, that confers its rhetorical force to persuade, what some authors call its 'narrativity' (Abbott, 2002).

However, it is important to recognize that this does not mean that faulty or incomplete causal inference is inherent in policy texts simply because of their narrative character. As argued by Megill (2007), such fallacies result from narrators' particular style of narration, where the narrative act lapses in argument and justification, and where historical, explanatory assertions are not properly backed up by supporting reasoning and facts. The persuasive power of such narration relies on various versions of the fallacy of *post hoc ergo propter hoc* (after this, therefore because of this), and it is normalized in policy texts through different means, a premier one being *ellipsis*, a form of simplification by omission, which we will explore further in the method section.

The unit of analysis in this study is impact arguments, particularly as they appear in impact evaluations of research and technology programmes, and how they conform to some of these rhetorical/narrative devices. This is the level of policy argumentation that Fischer (2003) refers to as 'technical-analytical discourse,' where the informal logic of the argument is addressed to considerations of fact. For programme evaluation there are (at least) three basic questions of verification on this level: Did the programme fulfill its objectives? Does the analysis uncover secondary, unanticipated effects that offset the programme objectives? Did the programme fulfill its objectives better than any alternative available means? (Fischer, 2003). Following the hermeneutic of suspicion, one might consider that a policy analyst 'working for' the programme principal would like to answer 'yes' to the first and third of the above questions, and no to the second. Wilson's (1973) (in)famous First Rule of policy evaluation states that this bias is realized by the commissioning agency suggesting the data, a time frame selected to maximize effects, and a programme contextualization that directs attention from alternative causes of those effects. Such active intervention, however, is rarely needed, since the

innate complexity and ambiguity of complex programmes lends every opportunity to practice what one may refer to as ‘narrative selectivity’ in considering the facts (*cf.* Hajer and Laws, 2006). In what follows, we will describe the material used and method applied in this study for uncovering such ‘narrative selectivity,’ and illustrate how it operates in impact evaluation reports.

Approach

Background and material

The case material for this study is drawn from three impact reports commissioned by the VINNOVA, the Swedish Government agency tasked with improving conditions for innovation in Sweden. The agency is located under the ministry for enterprise and innovation, and acts as Sweden’s link to the EU Framework Programme for R&D, and as an expert agency for innovation and growth related issues. The agency describes its activities as promoting ‘collaborations between companies, universities, research institutes and the public sector ... by stimulating a greater use of research, by making long-term investment in strong research and innovation milieus and by developing catalytic meeting places.’ (VINNOVA, 2017). This is commonly carried out by initiating and running research and technology programmes with a cooperative skew, typically involving Triple Helix constellations. These programmes are subsequently subject to impact evaluations (or ‘effect studies,’ to use VINNOVA’s chosen vocabulary), where the common denominator is to establish to what extent the overall agency goal (translated into more specific programme goals) of promoting sustainable growth, as well as specific goals pertaining to the programme, have been achieved. The evaluations usually take place a few years after the end of the programme in order for longer-term effects to have time to appear, and are always conducted by external evaluators.

The reports analyzed in this study were commissioned by VINNOVA to evaluate research programmes within two sector/industry specific areas: raw materials derived from renewable sources (Eriksson *et al.*, 2011) and innovations for future health (von Bahr, 2014). There was also a general regional development programme for supporting innovation in such areas as robotics, life sciences and bio-refineries (Kontigo, 2016). The first encompassed three sub-programmes which aimed at developing new commercially-viable products from renewables, thereby reducing environmental effects. The second programme was intended to support utilization of Swedish life science research in development of products, services and processes. The third programme aimed at sustainable regional growth by supporting and developing internationally-competitive research environments in the respective regions.

Analysis

The units of analysis for this study are not the evaluators or the case/programme context or the evaluation reports taken as whole arguments. Rather, they are the narrative statements found in the reports that express ‘impact arguments.’ Impact arguments are sequences of proposed premises and facts that are tied to conclusions about the impact of a programme or project. Such statements can be of various types and complexity, but

may include: (1) observations (naturally fallible and selective); (2) logical statements; (3) empirical statements on the relationship between observables; (4) methodological statements; (5) images and metaphors for integrating the above into stories; (6) value judgments about events and effects; and (7) normative policy conclusions arrived at by synthesizing the above (*cf.* Pen, 1985; Gasper, 1996). We are particularly interested in compound and singular statements encompassing 1–5 of the above, and in some cases 6 if the valuation component plays a role in the impact argument.

Our approach to the textual analysis has been to identify, categorize and analyze impact arguments, using a general inductive approach (Thomas, 2006). Such an approach sets out to identify units and commonalities in the text given a general interest, and then using such commonalities to derive conceptual patterns that can be used to order the phenomenon. In the present case, we were interested in capturing elements of simplification in impact arguments. The overarching principle of analysis, then, has been to capture what in literary analysis is known as ‘elements of ellipsis.’ The narrative device of ellipsis is enacted when the story (that which the text is about, the actual event sequence) is contracted in the narrated text, where the author ‘jumps over’ elements in the actual event sequence without accounting for them. Such jumps may be made explicitly (explicit ellipsis), as when the author indicates how much time has passed, or the nature of that which the jump covers, or implicitly (implicit ellipsis), as when no such indications are provided (Lothe, 2000). In the latter case, the reader might be left disoriented, not knowing how the author moved from A to B in his/her account, or of what the actual event sequence consisted. Of course, explicit and implicit ellipsis are matters of degree, and they may be more or less justified in a particular account, given narrative and thematic circumstances, such as previous explanations provided by the author, or the argumentative centrality of what was skipped over in the text.

There can be many reasons for ellipsis in a narrative about effects from some intervention. The omitted events may be perceived as unimportant to the general account. They may be perceived as self-evident, they may be unknown and therefore impossible to account for (but still their presence is assumed), or they may be difficult to put into words because of level of complexity and lack of knowledge (*cf.* Bal, 2002). Sometimes it is clear that the authors omitted central elements of the story; e.g., when they themselves indicate this is the case, when central elements of the sequence are summarized very briefly (a form of pseudo-ellipsis), or when there is a clear break in the account of a progression where causes simply do not add up to effects in the reader’s mind.

What counts as an ellipsis in the present study? We suggest that it is the absence of narrative-causal continuity in accounting for the progression from X–Y in the sense that it satisfies basic narrative conventions and expectations in the reader. To use a famous example from E.M. Forster (1927): ‘first the king died, then the queen died’ is not a model for a satisfactory impact statement. ‘First the king died, then the queen died from grief’ is a better candidate, since it adds the element of a conceptually-accessible causal pathway from X–Y. At this point, we note that there is no way of defining precisely and out of context when an impact argument becomes satisfactory. Good examples of satisfactory impact arguments have been strikingly difficult to acquire; however, we will end the results section by proposing

a (somewhat) successful example of such an argument, which can then be assessed in relation to previous, less-successful examples.

Results

We will symbolize the typical type of effect chain accounted for in impact evaluations by utilizing the convention where X represents the intervention/cause and Y represents the effects or outcomes of interest (e.g., Steel, 2008). In between intervention (X) and impact (Y) there is usually a number of mediating events or conditioning circumstances (A, B, C, etc.) that may be accounted for in more or less detail. The impact reports included in this study revealed three main types of inferential omissions, or ellipsis, in the causal chain from programme intervention to impact or effects. The first is the omission of all or some vital mediating events connecting X and Y. We refer to this as an ‘ABC ellipsis’. The second is where some ABCs are accounted for, together with an account of the impact Y, but an account of the intervention X has been omitted. We refer to this as an ‘X ellipsis’. The final one is where X and possibly an ABC chain is accounted for, but Y, the actual effect, is absent from the account, or has to be inferred. This is an example of a ‘Y ellipsis.’ The solid arrows in Figure 1 each denote one type of ellipsis. In what follows we will exemplify and explain each type, utilizing excerpts from the reports.

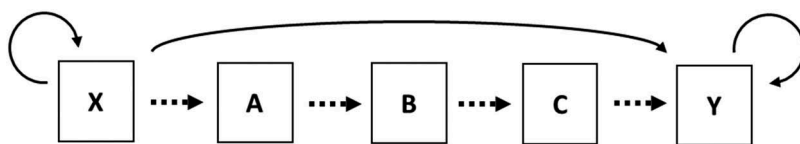


Figure 1. Three types of impact ellipsis (indicated by curved arrows).

ABC ellipsis

This is the most common form of simplification/omission found in the reports. In the case of ABC ellipsis, there is a reference to the intervention and then some form of jumping to conclusions, where key mediating events are omitted from the account. In the material, this was usually achieved by insinuating the link between X and Y rather than accounting for the connection by quoting spurious, specious links, or by speculating without empirical support. As an example of insinuation, we offer the following:

Several foreign companies have benefited from VINNOVA's support in various research projects. On one hand, Norwegian Borregaard, which is part of the research institute Inventia's business cluster, has built a pilot bio refinery, and also uses nanotechnology for food stuffs applications. In addition, Inventia has sold two patents to British firms. ... The patents related to compounds for treating skin cancer and hair loss respectively. (Eriksson *et al.*, 2011, p.35)

In this case, the authors quote the agency support as a cause of a number of effects, but omit reference to the actual connection or activities leading up to these outcomes. Instead, the connection is insinuated using the phrase ‘have benefited.’ The reader is left

with no clue as to how this benefit was brought about, other than the author's assurance. In other instances, the connection between funding and outcomes is inferred on the basis of spurious reasoning:

Almost all projects have been important in order to build up the researcher's and the companies' participation in national networks in the research area. This is a clear result of VINNOVA's work processes and rules, which encourage cooperation between different institutions and between firms. (Eriksson *et al.*, 2011, p.73)

In this case, the authors are observing an effect that should have been expected under normal circumstances, but assume a cause that is close at hand, perhaps because it was designated to be such a cause in the first place. This is a version of the *post hoc ergo propter hoc* fallacy. It is fallacious exactly because of a lack of a causal account of the ABC chain, and for not taking into account possible reasons for cooperation in the research field beyond the funder's work processes and rules.

The order of the account, as well as certain devices used to structure text, such as headings, can play a role in obscuring the connection between X and Y. In one of the reports, this was achieved by reversing the order of causality as it were, and beginning with a lengthy description of a result before accounting for actual influence of funding.

Results: Today there is a ready prototype that is functionality tested in relevant environments. The product itself, Scandivent, consists of two units, one that cleans the air and one that can be used to capture and diagnose airborne microorganisms/molecules. ... [the technical account continues over half a page] ...

Significance of the VINNOVA support: According to [NN] the funding was of great import in the initial phases of the project. Without the support the project group would not have achieved as much as they have today. (von Bahr, 2014, p.36)

In this case, the ABC connection is absent. However, the initial designation of a several paragraph-length technical account as 'results' suggests that this technology really was an outcome of the project funding. Now, when the reader reaches the actual description of the funding influence, this turns out to be something much less articulated. As a final example of this category, we offer an impact statement that, apart from being an ABC ellipsis, largely omits the X and the Y from the account as well:

[We see] how the carrying ideas and experiences that form the basis for the [x] programme also impact the region's work activities. Also in this case it is difficult to assert simple cause-effect relationships, or to pinpoint a direct effect of the [x] programme. It is rather that the [x] programme is part of a change in the way of thinking, which due to a clear profile and considerably long-term financial commitment could have an impact on a regional level. (Kontigo, 2016, p.25)

The causal agent here consists of 'ideas and experiences' of an indeterminate subject (the programme) which impact 'work activities' through some unknown pathways by dint of having a 'clear profile and long-term financial commitment.' The reader is given the impression of having been provided with a coherent impact argument, but closer scrutiny reveals its causal elements to be ambiguous and abstract.

X ellipsis

This form of omission can be referred to as an instance of ‘implying the cause’: some effects are accounted for, but description of relevant funding stimuli is wholly absent. A typical example is listing activities or outcomes associated with the participating actors:

We have identified 21 patents in total. This number is probably an underestimation given that some have reported that ‘several patents were taken’ without accounting for the exact number. . . . The number of prototypes may be somewhat underestimated in this study. We have identified a total of 19 prototypes. . . . In total we have identified eight products that have reached the market. . . . We have found five pilot plants. . . . We have identified seven new companies from the five different cases. (Eriksson *et al.*, 2011, pp.31–33)

Enumerations of effects without a clear account of the reason for them begs the question. This is partly because the relevant cause is already (supposedly) accounted for by the purpose of the report (namely to describe the effects of the funding), so that any event listed becomes an effect of that ‘cause.’ However, such tacit agreement between author and reader does not warrant listing ‘effects’ without any account of what caused them. A similar version is the following:

Results: Today the company produces seven different laminins. When the group applied for the VINNOVA grant they did not have any large scale production of any laminins, but that is the case today. . . . When the research group received the grant they had five customers, and today they have over 400 customers in 31 countries. (von Bahr, 2014, p.18)

Here again the focus is on effects, and the cause is simply implied. Similar to the third quote in the previous category (ABC ellipsis) the heading ‘results’ is used to intimate a causal connection between funding and outcomes. But the reader is not given any indication of how these outcomes were actually effected. X ellipsis can also be achieved by quoting an action as being performed by ‘the funded project’ as such, and thereby implicitly asserting that the funding is the cause of that action, although of course it was performed by individuals, not a project:

Also in the development of the Strategic Innovation Agendas [another government programme] we observe how several [x] initiatives have had an important role in initiating and mobilizing around the development of the Agenda and in implementing [the programme]. (Kontigo, 2016, p.24)

Our final example of an X-ellipsis is where the author first challenges the attribution of cause to the funding, where some subsequent action is concerned, only to later assert or imply it without any further explanation:

We cannot clearly say that VINNOVA has been instrumental in this change. [Rather there is] a general development, in industry and policy, towards environmental concerns as an area for innovation. However, we propose that it is clearly so, that a large part of the innovative work conducted within the [x] programme has used environmental efficiency as a tool for furthering innovation within this growth area. (Kontigo, 2016, p.34)

The key element in this quote is the word ‘however,’ which denotes that in spite of environmental concerns being a general trend, there is still reason to believe that programme funding was responsible for them in this instance. The reasons for this are omitted from the account.

Y ellipsis

A Y ellipsis occurs when there is some general effect claim, but the nature of this effect is conjectured, or is so vague as to be completely indeterminable:

It is very common that VINNOVA projects lead to follow on projects. Participating in a network and thereby increasing the exchange of knowledge broadens the vision between researcher and departments. One effect of this could be that new combinations of knowledge and experiences give rise to new innovative ideas that could generate new research projects. At least 36 of the projects studied have led to one or several follow-on projects. (Eriksson *et al.*, 2011, p.36)

In this case, the ‘effect’ seems to amount to a number of hypothesized, positive outcomes of follow-on projects which supposedly stimulate networking etc. The chain of events that should follow from such projects is overtly speculative, yet it is claimed as an effect by associating 36 funded projects with this general category. We are none the wiser as to whether these 36 projects actually experienced these networking effects. While this is a case of conjectured effect, the following demonstrates vagueness in addition to conjecture:

Significance of the VINNOVA support: The grant made a big difference according to [NN], since without it the project would have probably been shut down. After VINNOVA’s grant had been used up, there was another grant from [council] that enabled further development of the project idea. (von Bahr, 2014, p.16)

Here the effect of the grant seems simply to be that it kept the project running, but there is no description of what actually resulted. Beyond a spurious reference to another council having picked up the project after the grant finished, the effect account seems completely circular – the effect of the funding was that the project got funded. Effect claims can also be so elusive that they are impossible to assess or verify qua effects of anything:

In Västernorrland [region] there is today, partly through [project]’s work, a significantly better anchored understanding of what it means to work according to a Triple Helix perspective when it comes to developing the bio-economy. A clear effect of this is that the region has an interest in mustering resources towards yet another investment [project]. (Kontigo, 2016, p.40)

We are offered two effects, which are equally elusive. The first is ‘a better understanding’ of what it means to work according to ‘a Triple-Helix perspective,’ and the other is the emergence of ‘an interest’ to pursue another project. The place where the two effects supposedly materialized is simply ‘the region,’ which makes the target as elusive as the proposed effects.

Our final contribution to this section is a counterpoint to the above examples, namely an impact argument that, in our opinion, is not an oversimplification in the sense of an unjustified ellipsis. This means that the impact argument as such, while omitting causal intermediaries, still includes causal elements offering the reader a viable pathway from cause to effect that is narratively satisfactory:

[Project X] has since its inception in 2004 improved the strength of its brand and established better links with other actors in the regional innovation system. This has meant that [X] has gained a better insight into the other actors’ development strategies, and has been able to influence local actors to increase their priority on IT and automation in their research and innovation. ... This position made it possible for [X] to drive the process to develop the strategic innovation programme [Y]. (Kontigo, 2016, p.40)

This impact argument is, of course, far from perfect. There is a tendency towards X ellipsis, and the outcome is a bit obscure ('develop the programme ...'). However, it is good enough to indicate a better way to construct these statements. There is a beginning, middle and an end, and the mediating events (ABC) make narrative sense in terms of how they appear to enable the outcomes. There is enough causal detail to make the connections referred to convincing. With this positive example of an impact argument, we will now offer some concluding reflections on what this study might offer in terms of general insight into thinking about impact evaluations in R&I.

Discussion and conclusions

The quotes above provide some examples of unsatisfactory and misleading impact arguments. Apart from the fact that they seem to pertain to different parts of the narrative sequence of beginning (cause), middle (intermediating factors) and end (effects), these examples demonstrate certain forms of simplification. Under the ABC ellipsis, we typically observe instances of what may be referred to as 'omission'. This is where the connection between funding and impacts is insinuated by using a specific phrase (such as 'have benefited'), which is then not explained. Connections are also inferred on the basis of spurious reasoning (such as when an active cause from funding is assumed, but where the outcome could have been expected under normal circumstances). The ABC ellipsis also illustrates the rhetorical principle of structuring. This is where temporal order in the argument is used to replace an account of a chain of events (as in the *post hoc ergo propter hoc* fallacy). Here the author uses narrative order and textual structure to achieve a sense of impact; for example, with lengthy descriptions of 'results', followed by considerably shorter and less explanatory 'significance of support' accounts. The rhetorical element is also clearly present in the final ABC account, which we like to term an instance of 'abstraction'. This is where ambiguous and abstract narrative elements assume causal significance: the relationships described are actually conceptual, and not at all accounted for in actual causal terms.

The other two forms of ellipsis offer examples of the same patterns of simplification. The X ellipsis, for example, demonstrates an instance of simplification, where effects are enumerated without visible cause: a question begging omission where the reader is expected to assume that funding created these effects. As in the ABC examples, we also observe the structuring use of headings ('results') to intimate causal significance, but without any account of what actually produced these results. The rhetorical ploy is that 'results' are always the result of something; why spell out this something when it should be obvious? As with the ABC examples, there is also a case of abstraction here in the designation of the project as the agent of change rather than a tangible actor capable of doing anything. Both the X and the Y ellipsis include examples of overt omission: cause and effect accounts are completely hypothesized or asserted for no reason whatsoever. For instance, in the final example of X ellipsis, the authors provide no grounds for inferring a funding cause, but simply assert it. The Y ellipsis offers more examples of hypothesized or speculative outcomes; for example, networking effects that are assumed but not demonstrated. Yet another example of omission (in the sense of spuriousness) is where impacts are just a first order effect of the funding itself, as in the empty or tautological inference that, because of the funding, a project 'was able to continue' (and then what?).

Stated simply, the present study presents a type of analysis that we believe ought to be performed in the course of conducting impact evaluations and reporting on (perceived) impacts. Taking the pitfalls and fallacies exemplified above into account can raise sensitivity to certain analytical traps and offer methodological insight into the general challenges in accounting for impact and change in research and technology processes. There are several issues involved in producing locally and generally valid accounts of technological change. Local accounts suffer from the problems of attribution discussed above, specifically related to challenges of the counterfactual (e.g., Gertler *et al.*, 2016). Eliciting counterfactually valid impact accounts from involved stakeholders requires more than just asking them what impact the programme had. There are problems of hindsight bias and lack of overview and memory (Granhag *et al.*, 2000).

The need to generalize about impacts e.g. in producing programme theory (logic models), represents another set of challenges (*cf.* Penfield *et al.*, 2014). These are connected to inherently tricky aspects involved in the historiography of innovation (see Popper, 1963). They pertain to what can be called ‘the inexhaustibility of description’ of knowledge-based change. There is a need to select aspects of history believed to be relevant for theory (explanation and prediction). However, a selection can never be exhaustive, or even trusted to be representative. This is because development in research and technology is dependent on essentially unpredictable changes in knowledge (*pace* Popper), and also on long-term cumulative factors and unpredictable path-dependencies, complementary resources and network effects, and the effects of generic technologies in releasing the potential of existing knowledge (e.g., Girfalco, 1991).

There is no avoiding simplification in accounting for these processes and anyway, simplification is one basic goal of science. What we are interested in here involves the justification of descriptive and explanatory simplifications. White (2010) suggests that an impact evaluation should at some minimum connect the activities of a programme to its goals and effects. Majone (1989) argues that the standards we employ in choosing to accept certain simplifications (and the level of rhetoric of these accounts) represents an important *sine qua non* for policy analysis. One might assume that the acceptability of such accounts rests on narrative principles, including *prima facie* assessment of reasonableness in the way a reader perceives impact arguments. The trend in evaluation research seems to be to remedy these inferential challenges either by theorizing change processes *ex ante* or *ex post*, or by using control group designs (in which case one does not need a theory of change). Another approach is a more variegated attitude to the challenge of making causal claims in evaluations. Following Gates and Dyson (2017), this involves not only producing relevant and defensible causal arguments, but also being responsive to the contextual aspects of the intervention and the situation; addressing inferences from the point of view of multiple audiences, demonstrating sensitivity (familiarity) with several ways of thinking about causality, designs and methods, and recognizing that theory and causality operate on several (layered) levels.

What this article provides is not a picture of determinable, acceptable limits for narrative simplification in impact arguments and evaluations. Rather, it is an account of a number of narrative fallacies that, if not controlled, may lead to unsatisfactory explanations of impacts, in the sense that they leave too many parts of the effect chain implicit, and do not provide narratively-acceptable justifications (logically and empirically) for their conclusions. We

have managed to capture some of the devices involved in what we believe is unjustified ellipsis: direct omission of factors (for instance, in the form of insinuations of their presence, spurious inferences, hypothesized, tautological or speculative outcomes); the use of abstraction to forge the impact argument (such as transferring agency to conceptual categories or entities); and the use of structuring of the impact argument (for example, achieving a sense of impact through the narrative order the account is given and by the use of headings). We have offered a few suggestions as to what principles might be applied to forge more complete and reflexive impact arguments: we leave it to the reader to judge how these considerations might be used.

Acknowledgments

This work has been funded by the Swedish Foundation for Humanities and Social Sciences.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Abbott, H. (2002) *Narrative*, Cambridge University Press, Cambridge.
- Bal, M. (2002) *Narratology: Introduction to the Theory of Narrative*, University of Toronto Press, Toronto.
- Cartwright, N. (2007) *Hunting Causes and Using Them: Approaches in Philosophy and Economics*, Cambridge University Press, Cambridge.
- Cartwright, N. and Hardie, J. (2012) *Evidence-Based Policy*, Oxford University Press, New York.
- Chatman, S. (1978) *Story and Discourse: Narrative Structure in Fiction and Film*, Cornell University Press, Ithaca NY.
- Donovan, C. (2011) 'State of the art in assessing research impact. Introduction to a special issue,' *Research Evaluation*, 20, pp. 175–79. doi:10.3152/095820211X13118583635918.
- Donovan, C. and Hanney, S. (2011) 'The 'Payback Framework' explained,' *Research Evaluation*, 20, pp. 181–83. doi:10.3152/095820211X13118583635756.
- Eriksson, T., Roback, K., Lundqvist, M., Levin, L. and Carsson, P. (2011) *Effektanalys Av Forskningsprogramme Inom Material Från Förnyelsebara Råvaror [Effect Analysis of Research Programmes on Materials from Renewables]*, VINNOVA, Stockholm.
- Fischer, F. (2003) *Reframing Public Policy: Discursive Politics and Deliberative Practices*, Oxford University Press, Oxford.
- Fischer, F. and Forrester, J. (1993). *The Argumentative Turn in Policy Analysis and Planning*, Duke University Press, Durham NC.
- Fischer, F. and Gottweis, H. (eds) (2012) *The Argumentative Turn Revisited: Public Policy as Communicative Practice*, Duke University Press, Durham NC.
- Forster, E.M. (1927) *Aspects of the Novel*, Harcourt, Brace & Co, New York.
- Gasper, D. (1996) 'Analysing policy arguments,' *European Journal of Development Research*, 8, pp. 36–62. doi:10.1080/09578819608426652.
- Gates, E. and Dyson, L. (2017) 'Implications of the changing conversation about causality for evaluators,' *American Journal of Evaluation*, 38, pp. 29–46. doi:10.1177/1098214016644068.
- Gertler, P., Martinez, S., Premand, P., Rawlings, L. and Vermeersch, C. (2016) *Impact Evaluation in Practice, Second*, Inter-American Development Bank and World Bank, Washington DC.
- Girfalco, L. (1991) *Dynamics of Technological Change*, Van Norstrand Reinhold, New York.

- Gläser, J. and Laudel, G. (2015) 'A bibliometric reconstruction of research trails for qualitative investigations of scientific innovations,' *Historical Social Research*, 40, pp. 299–330.
- Granhag, P., Strömwall, L. and Allwood, C.-A. (2000) 'Effects of reiteration, hindsight bias, and memory on realism in eyewitness confidence,' *Applied Cognitive Psychology*, 14, pp. 397–420. doi:10.1002/1099-0720(200009)14:5<397::AID-ACP658>3.0.CO;2-M.
- Guba, E.G. and Lincoln, Y.S. (1989) *Fourth-Generation Evaluation*, Sage Publications, Thousand Oaks CA.
- Hajer, M. and Laws, D. (2006) 'Ordering through discourse' in Moran, M., Rein, M. and Goodin, R. (eds) *Oxford Handbook of Public Policy*, Oxford University Press, Oxford, pp. 251–68.
- Hogwood, B. and Gunn, L. (1984) *Policy Analysis for the Real World*, Oxford University Press, New York.
- Kontigo. (2016) *Effektanalys Av Vinnväxt-Programmet: Analys Av Effekter Och Nyttä [Effect Analysis of the Vinnväxt-Programme: Analysis of Effects and Utility]*, VINNOVA, Stockholm.
- Kostoff, R. (1993) 'Semiquantitative methods for research impact assessment,' *Technological Forecasting and Social Change*, 44, pp. 231–44. doi:10.1016/0040-1625(93)90070-N.
- Lothe, J. (2000) *Narrative in Fiction and Film*, Oxford University Press, Oxford.
- Majone, G. (1989) *Evidence, Argument and Persuasion in the Policy Process*, Yale University Press, New Haven CT.
- Mansfield, E. (1998) 'Academic research and industrial innovation: an update of empirical findings,' *Research Policy*, 26, pp. 773–76. doi:10.1016/S0048-7333(97)00043-7.
- Megill, A. (2007) *Historical Knowledge, Historical Error: A Contemporary Guide to Practice*, University of Chicago Press, Chicago.
- Mohr, L. (1995) *Impact Analysis for Programme Evaluation*, Sage, Thousand Oaks CA.
- National Academy of Sciences. (1999) *Evaluating Federal Research Programmes: Research and the Government Performance and Results Act*, National Academies' Press, Washington DC.
- O'Keefe, C. and Head, R. (2011) 'Application of logic models in a large scientific research programme,' *Evaluation and Programme Planning*, 34, pp. 174–84. doi:10.1016/j.evalprogplan.2011.02.008.
- Pawson, R. and Tilly, N. (1997) *Realistic Evaluation*, Sage Publications, London.
- Pen, J. (1985) *Among Economists*, North Holland, Amsterdam.
- Penfield, T., Baker, M., Scoble, R. and Wykes, M. (2014) 'Assessment, evaluations, and definitions of research impacts: a review,' *Research Evaluation*, 23, pp. 21–32. doi:10.1093/reseval/rvt021.
- Popper, K.R. (1963) *Conjectures and Refutations: The Growth of Scientific Knowledge*, Routledge, London.
- REF (2014, 02) *Assessment framework and guidance on submissions*, available from <http://www.ref.ac.uk/pubs/2011-02/> [accessed Jun 2017].
- Roe, E. (1994) *Narrative Policy Analysis: Theory and Practice*, Duke University Press, Durham NC.
- Scriven, M. (2008) 'A summative evaluation of RCT methodology: & an alternative approach to causal research,' *Journal of MultiDisciplinary Evaluation*, 5, pp. 11–24.
- Smith, K. and Larimer, C. (2017) *The Public Policy Theory Primer*, Westview Press, Boulder CO.
- Steel, D. (2008) *Across the Boundaries: Extrapolation in Biology and Social Science*, Oxford University Press, New York.
- Stern, E. (2013) 'Editorial,' *Evaluation*, 19, pp. 3–4. doi:10.1177/1356389012473814.
- Thomas, D. (2006) 'A general inductive approach for analyzing qualitative evaluation data,' *American Journal of Evaluation*, 27, pp. 237–46. doi:10.1177/1098214005283748.
- VINNOVA (2017) available from www.vinnova.se [accessed Jun 2017].
- von Bahr, J. (ed). (2014) *Hälsoekonomisk Effektanalys Av Forskning Inom Programmet Innovationer För Framtidens Hälsa. [Health Economic Effect Analysis of Research within the Programme 'Innovation for Future Health']*, VINNOVA, Stockholm.
- Weiss, C.H. (1998) 'Have we learned anything new about the use of evaluation?' *American Journal of Evaluation*, 19, pp. 21–33.
- White, M. (2010) 'A contribution to current debates in impact evaluation,' *Evaluation*, 16, pp. 153–64. doi:10.1177/1356389010361562.
- Wilson, J. (1973) *Political Organizations*, Basic Books, New York.

Wooding, S., Hanney, S., Buxton, M. and Grant, J. (2005) 'Payback arising from research funding: evaluation of the arthritis research campaign,' *Rheumatology*, 44, pp. 1145–56. doi:[10.1093/rheumatology/keh708](https://doi.org/10.1093/rheumatology/keh708).