The unused potential of process tracing as evaluation approach: The case of cluster policy evaluation

Michael Rothgang and Bernhard Lageman
RWI – Leibniz Institute for Economic Research, Germany

Abstract
This article shows that process tracing developed in social science research can be used in evaluations of complex structural and technology policy programmes to overcome deficits in the methodological instruments used to date. Cluster policies are a well-suited example because they are characterized by complex impact patterns like many other current structural and innovation policy programmes. The origin and characteristics of the methodological approach of process tracing are discussed and weaknesses of impact evaluations of cluster programmes highlighted. Subsequently, we look at the potentials of process tracing in impact evaluations of cluster programmes within the framework of mixed-method designs. Our analysis shows that process tracing can enrich the applied methodological repertoire. It allows the evaluators to test the accuracy of the theoretical assumptions underlying the analysed programme and to identify causal mechanisms.

Keywords
causality, cluster policy evaluation, mixed method, policy impact, process tracing

Introduction
The starting point for our analysis is the observation that there is a gap between the methodological equipment of evaluators and the questions that are addressed by evaluations of public interventions. This applies especially to public interventions of a complex, systemic nature like in the health sector, regional development or innovation policy. When programmes are characterized by multi-dimensional target systems and the complex causal mechanisms
cannot easily be grasped in statistical models, mixed-methods evaluation designs are useful to understand causal relationships. These evaluation designs combine quantitative and qualitative elements on an equal basis (triangulation).

One rather recent methodological approach, which has the potential to strengthen the explanatory power of qualitative research and to build bridges between quantitative and qualitative approaches, is *process tracing* (PT). Until a decade ago, PT has been applied in several social sciences such as political science, sociology, psychology, and historical or educational science. Only recently, it has found increased attention in the evaluation of public policy programmes, particularly in the context of international development projects (see, for example, Barnett and Munslow, 2014; Befani and Stedman-Bryce, 2017). But PT has hardly been used in the evaluation of cluster and innovation policy support programmes (exceptions are Bjurulf et al., 2012; Bramwell et al., 2019). Against this backdrop, our contribution deals with the unused potential of PT in cluster policy impact evaluations.

Cluster policy presents itself as a suitable field for the application of PT for several reasons. Cluster policy has attracted a lot of attention in recent decades. Under cluster policy (or cluster policies), we subsume policies that promote the development of local clusters of firms, research organizations and other organizational entities. Their aim is to increase the competitiveness of the cluster actors themselves, of the cluster of cooperating actors in the region (firms, research organizations and other actors) and of the regional economy. Although the peak of public attention for cluster policy seems to have passed, it will certainly remain an important instrument of structural policy (Wilson, 2019). By their very nature, cluster programmes are often rather complex in respect to their impact patterns. They follow target systems that cannot be reduced to one or only a few success parameters. Therefore, the application of the available methodological evaluation tools has so far led to rather fragmented insights regarding the impact of cluster policy interventions (Kiese, 2017; Lindquist et al., 2013; Rothgang and Lageman, 2021; Uyarra and Ramlogan, 2016, 2017). For this reason, we see a possibility to broaden the range of evaluation methods by exploiting the potential of PT.

This article deals with the possibilities and preconditions of the application of PT in cluster-related impact evaluations. We investigate whether and to what extent the PT approach would be suitable to fill an obvious gap in the methodological repertoire of impact evaluation. There are other possible approaches under the umbrella of theory-based evaluation (TBE) that are not explicitly contrasted to PT but addressed in the section ‘PT: Concept and application in evaluation contexts’.

Our article is structured as follows: the ‘PT: Concept and application in evaluation contexts’ section deals with the concept of PT and its application in evaluation contexts. The ‘Missing links in the cluster evaluation methodology repertoire’ section discusses in what respect PT is suited to fill a lacuna in the methodological toolkit of cluster policy impact evaluations. The section ‘Developing PT to be a useful method in cluster evaluation’ develops concrete suggestions for the application of PT in cluster policy impact evaluation. The ‘Summary and discussion’ section summarizes the results and discusses implications for further research.

**PT: Concept and application in evaluation contexts**

The use of PT in qualitative social research as a recognized approach following rules shared by a broader research community is quite recent. However, the core idea of this approach dates
back to the beginnings of historical research in antiquity (for the European tradition Thukydides, cf. Bennett and Checkel, 2015b). Although these predecessors did not explicitly claim to argue for case-based causal inferences on an epistemologically sophisticated basis, it is important to note that there was causally oriented PT *avant la lettre*. The recent ‘reinvention’ of PT and the idea to develop a clearly formulated methodological basis for this method originates from individual psychology. Later, the method was widely used not only in political science (e.g. the contributions in Bennett and Checkel, 2015a) but also in other disciplines such as sociology and criminology. Remarkably, PT as an explicit methodological approach has so far hardly been applied in economic studies.

A literature search confirms, on one hand, that PT as an explicit approach of investigation is a relatively new phenomenon and, on the other hand, that the interest in this subject has recently increased strongly. A general search (no query restrictions) in the literature database Scopus from 27 November 2020 identified 5631 documents registered until 2019 that contain the term ‘process tracing’. The number of publications is reduced substantially when we regard only papers in which PT plays an important role in the analysis. If the search was reduced to the three items ‘article title’, ‘abstract’ and ‘keywords’, this number decreased to 1104. Over time, a remarkable growth can be observed with the numbers increasing from 21 (10 for search in title, abstract or keywords) in the decade 1970–1979 to 4034 (790) in 2010–2019, the decade with the by far highest publication figures to date.

As Collier (2011) points out, the development of PT is part of a larger effort to sharpen and systematize the methods of qualitative research (e.g. Brady and Collier, 2004; Gerring, 2008, 2019), which took place parallel to the significant innovations of quantitative research tools for causal analysis (e.g. the rise of randomized controlled trials to become the gold standard for the analysis of causal relations). Moreover, the ongoing discussion about the ‘replication crisis’ in medical and social science research has created a greater sensitivity for qualitative procedures and methodological triangulation. As PT has gained acceptance in various research contexts, the interpretations of this approach have also become more diverse. This explains the marked differences between existing definitions. In the following, we first deliver our definition of PT and then discuss the epistemological basics, the variants and the possible application of PT in impact analyses of complex programmes. Our definition of PT ties in with the works of Bennett and Checkel (2015a, 2015b), Beach and Pedersen (2019), and Collier (2011) and takes its own position on the points in dispute within the community of researchers using this method.

In our context, we understand PT as an analytical approach which aims to investigate the causal mechanisms (for the concept, cf. Bennett and Checkel, 2015a, 2015b; Gerring, 2008; Hedström and Ylikoski, 2010) that have led to a particular event or to an observed development of an object of investigation. The period under consideration ranges from an initial to a defined final state. We prefer the term ‘approach’ to characterize PT, because the analysis can be based on a wide range of information sources (documentary sources, primary or secondary statistics, and professional studies) and methodological instruments of data collection and analysis (standardized surveys, interviews, statistical investigations and model analyses). Consequently, there is not a single clearly defined method or procedure to be equated with PT. Rather, this approach is based on the pragmatic use of one or several methods, whereby the choice of the concrete method(s) depends on the object of investigation and the purpose of the study.
Central to PT is the decoding of the causal mechanisms underlying the development of the object of investigation. This claim implies an epistemological position based on scientific realism (Abbott, 2001; Bennett and Checkel, 2015a, 2015b). The challenge is to determine, within the framework of a meticulous process-oriented analysis of a single case, whether a presumed causal relationship between influencing factors and observed outcomes in the development of an object of investigation between an initial and a final stage can be confirmed on the basis of the identified facts or whether a corresponding initial causal hypothesis proves to be false. In case the theoretical starting hypothesis is refuted, the question arises whether other explanatory factors (a competing theory) are available for the explanation of the occurrence of the observed result.

The research approaches carried out under the label ‘process tracing’ are manifold and – apart from the shared basic principle – can hardly be reduced to a common denominator. This caused Beach and Pedersen (2019) to identify different variants of PT depending on their respective function in the knowledge process. In the second edition of their monograph, they distinguish four types of PT that aim at different aspects of investigation (p. 9):

(i) **Theory-testing PT** checks whether a previously hypothetical assumed causal mechanism is present and works as assumed.

(ii) **Theory-building PT** is aimed at finding out which causal mechanisms led to the final state without having defined it in advance.

(iii) **Theoretical-revision PT**, like theory-testing PT, starts from an ex ante given theory and asks why the presumed causal mechanism was not present.

(iv) **Explaining-outcome PT** looks for an adequate causal explanation of an event that has occurred.

While the first three approaches address the applicability of a theory in a concrete situation or the discovery of an anomaly, the latter approach focuses on the understanding of a concrete case. In most of the so far existing applications of type (iv), the analysis focuses on the explanation of a historical event.

None of the four cases by itself fully describes the specific research constellation we are addressing in this article, namely, the evaluation of the impact of complex government programmes. For impact evaluations, all four types of PT have to be applied. **First**, PT uses the ‘theory’ or theoretical assumptions on which the programme is based as its obvious starting point. By means of a precise examination of the processes that have taken place, the evaluators scrutinize, whether the assumed causal relationships can be confirmed in the specific case or must be rejected. If discrepancies are found, they ask whether there are deficits in the formulation of the programme theory. **Second**, other causal mechanisms which also influenced the development of the object of investigation are identified and their impact on the overall result is analysed. **Third**, obvious deviations of the observed from the assumed causal mechanism open up the opportunity to question the theoretical basis of the policy intervention, to rethink, and, if necessary, revise the promoted causal relationship. **Fourth**, PT in the context of an impact evaluation is case-centred and looks at the causes leading to a defined final state of the investigation object.
In the generic context of the evaluation literature, the concept of PT can be included in the broad circle of evaluation procedures discussed under the heading of ‘theory-based evaluation’ or, in a different choice of words, theory-driven or theory-oriented evaluation (for the general concept, cf. Centre of Excellence for Evaluation Government of Canada (CEE), 2012; Coryn et al., 2011; Giel, 2013; Pawson, 2013). The evaluation approaches classified under this term have in common that they are used to analyse the effects of a programme based on the specifications of the programme theory which in most cases is only implicitly communicated by programme makers. In such cases, the theory has to be analytically worked out by the evaluators. The programme theory describes the expected relationships between programme impulse, impact mechanism and programme objectives.

In the past decades, TBE has been widely adapted in evaluation research and has become increasingly differentiated into several approaches. Today, there are hardly any programme evaluations that do not refer, at least in a rudimentary form, to the underlying programme theory. However, this also opens the door to a certain arbitrariness in terms of the methods used to analyse the implementation of the programme’s objectives and the adequacy of the programme theory. This prompts us to focus on PT in this article, as it addresses a crucial weakness of the methodological toolkit usually applied in cluster policy impact evaluations today.

The approach proposed here is closely related to contribution analysis, an approach that focuses on the specific contribution of a programme to the development of the object of study, separating the effects of the programme impulse of those of other factors (Mayne, 2012). The procedures chosen by evaluators under contribution analysis are very similar to PT with regard to their analytical intention and their typically flexible choice of methodological instruments (American Evaluation Association, 2015; Befani and Stedman-Bryce, 2017).

The evaluation literature also sometimes refers to ‘process analyses’ in a sense, which is clearly different from PT. In these instances, this concept is understood to mean the analysis of the implementation process of a programme in terms of its correct management, the spending of funds, the role of project agencies, obstacles in its implementation and the like (e.g. American Evaluation Association, 2015; Gertler et al., 2016). Other authors understand process analyses to be largely identical to PT (e.g., Grant et al., 2013). Like PT, process analyses in this second understanding of the approach have so far played only a marginal role in the field of cluster policy impact evaluation.

**Missing links in the cluster evaluation methodology repertoire**

Research on clusters has produced a substantial amount of empirical and theoretical studies from different disciplines. At the same time, cluster policy in the developed world has experienced a notable upswing since the mid-1990s, receiving decisive impulses from cluster research. It is therefore surprising that the evidence for effects of cluster policy is still rather ambiguous. There are some indications of positive economic, technological and social impacts of evaluated cluster programmes, but also of programmes that obviously failed (Andersson et al., 2004; Fornahl et al., 2015; Fromhold-Eisebith and Eisebith, 2008; Kiese, 2017; Lindquist et al., 2013; Uyarra and Ramlogan, 2016, 2017).

A review of journal articles and commissioned evaluation studies dealing with the impact of cluster policy on the development of funded clusters in Germany (Rothgang and Lageman, 2021) concluded that systematic information on the long-term development of promoted
clusters is virtually absent. Information on promotion effects is limited to a narrow circle of more easily accessible effects. The analysis comes to the conclusion that an assessment of the contribution of public funding to cluster development is hardly possible on the basis of the available study results. Regional development contexts, which could provide information about the impact of competing factors, are not systematically analysed.

Several causes can be cited for the current unsatisfactory state of knowledge about the effectiveness and efficiency of cluster policy: institutional circumstances, such as the predominance of contract research and the lack of interest on the part of public funding bodies in financing large academic evaluation projects, play a role. The long periods of time required for the effects of complex programmes to unfold make impact evaluations difficult and time-consuming. The interest of policy makers is primarily focused on current political events, so the effectiveness of programmes implemented by predecessor governments is often of little importance to them. Furthermore, the difficult accessibility of relevant data is a factor. Finally, despite the progress made in recent decades, the methodological tools applied in impact evaluations do not provide fully satisfactory solutions to three key methodological challenges (see also Edler and Fagerberg, 2017):

- **Attribution problem**: Clusters are complex entities that are simultaneously exposed to many external and internal influences in their development, including emergent processes. Since different causal mechanisms are at work and interplay with each other, evaluators face the challenge of attributing the results of cluster developments to various causal mechanisms.
- **Context dependence**: Clusters when being promoted always carry strong individual, idiosyncratic features. The particular context in which the support is provided is of utmost importance for the success of cluster promotion. Methodological procedures tailored to individual cases should be able to uncover the context and show under which conditions the guiding principle of the programme works.
- **Systemic character**: Cluster promotion aims at an object which is of a complex nature and can only be adequately identified through systemic analysis. The common research methods mostly fail to reflect the systemic nature of the phenomenon.

At first glance, it may seem strange to introduce a processual approach into a field of research that essentially lives from the analysis of development processes. So, it seems necessary to ask, whether the careful analysis of the causes and mechanisms of cluster development is to some extent ‘built into the genes’ of impact evaluation.

Generally, cluster research has dealt extensively with descriptive analyses of the development of individual clusters or, on a comparative basis, the development of industrial clusters in larger economic areas. Porter already based his theoretical approaches on comparative empirical analyses of development processes in clusters, regions and nations (Porter, 1998). Following Porter and earlier research on economic agglomeration processes, a large number of scientific publications have examined the development of clusters from an economic perspective since the mid-1990s. These studies had their forerunners in the analyses of Marshall’s industrial districts or of the Italian *distretti industriali*. As a result, a wealth of case studies examines the development of individual clusters and provides valuable approaches for theory building (e.g. the examples in Karlsson, 2008). An example of studies with more theoretical claims is the literature on cluster life cycles which has given fruitful impulses to further
research not only on clusters but also on cluster policy approaches (e.g. Bergman, 2008). At the same time, authors who contributed to the fields of regional economics and evolutionary economic geography have dealt extensively with regularities in the development processes of firm clusters and urban agglomerations (Bathelt and Glückler, 2018; Krugman, 1997).

The main focus of cluster policy evaluators is the influence of a promotion programme on the development of a cluster within a defined time frame (e.g. Fornahl et al., 2015). In cluster policy impact analyses, development processes of promoted clusters are therefore inevitably objects of the investigation. This does not mean, however, that in practical evaluation studies, the cluster development processes would be systematically examined with regard to their underlying multiple causalities. Qualitative studies usually do not include a detailed analysis of the processes taking place (e.g. Dohse, 2000). Econometric analyses like, for example, quasi-experimental approaches also aim to capture the results of developments that have occurred over time and are in this sense process-oriented (e.g. Lehmann and Menter, 2017). However, they are only casually interested in particular impact mechanisms and only lead to valid results if the underlying data and approach are sufficiently robust to balance the influence of intervening variables. Therefore, the application of econometric methods is restricted to a few variables like the effects on R&D expenditure or the number of product/process innovations. In addition, the multitude of factors that influence outcomes makes it advisable to check the results of econometric calculations by other methods.

In short, although causal mechanisms determining development processes of clusters are a central object of interest in cluster research, they are only marginally considered in impact evaluations of cluster policy. The main reason for this is that the methodological toolkit is not focused on detailed analyses of policy impact chains in cluster development and the exploration of competing causal mechanisms in complex programme constellations.

**Developing PT to be a useful method in cluster evaluation**

In the following, we address four core aspects that are important to develop PT into a useful method in cluster impact evaluation: the rationale of the application of PT, the role of theory, the possible objects of investigation and the role of PT as a component of evaluation designs.

**The rationale of PT in cluster impact evaluation**

Impact evaluations usually focus on comparing the final state with the initial state of the funded clusters with the help of, for example, a statistical model or a retrospective survey of key actors in order to determine whether or to what extent the promotion has achieved the desired purposes (Figure 1).

However, there is usually no detailed assessment of the processes occurring between the initial and final state. These are rather treated as black boxes. This is where PT can make a substantial complementary contribution to the understanding of policy impact. The relevant processes are subsequently reconstructed or analysed in an accompanying or ex post evaluation. PT can include among other factors, the analysis of the decisions and actions of the relevant actors; the factors influencing the development of cluster networks, companies and research organizations; the processes of knowledge accumulation; and the influences of the economic and institutional environment.

The task of PT is to examine the processes taking place in the promoted cluster between the start of the support and the point of reference for determining the result. This should be done
systematically and in accordance with the principles of PT described in the methodological literature (e.g. Beach and Pedersen, 2019; Bennett and Checkel, 2015b; Collier, 2011) and the general methodological guidelines of the associations of evaluation researchers (e.g. DeGEval, 2017). In more complex programmes, the funding will usually include several clusters. The research task is to analyse which causal mechanisms are responsible for the observed developments (Figure 2). From the perspective of the client of the evaluation, and eventually of the evaluators, there are two main questions that arise: On one hand, what effects result from the original programme impulse? On the other hand, can the guiding theory underlying the programme’s design be confirmed/not confirmed or does it need adaption? Besides these questions, any other causal mechanisms influencing the final state of the cluster are analysed.

These might make further adjustments regarding the theoretical framework necessary.

From an epistemological point of view, neither the confirmation nor the refutation of the underlying theory can be taken as a final verdict on its applicability. Rather, the confirmation of the theory underlying the intervention in the current case may provide a good argument for the design of similar interventions in the future. On the other hand if the theoretical argument proves to be (partly) inaccurate, there would be an anomaly in need of explanation. If the result is ambiguous, at least some information would be provided on the conditions under which the intervention could be more successful.

**The role of theory in PT-based impact evaluation**

The analyses of cluster development processes mentioned in the ‘Missing links in the cluster evaluation methodology repertoire’ section either fulfil a primarily descriptive function or
they serve the goal of developing elements of a theory of cluster development. In contrast, the task of PT in impact evaluations is to check whether the causal assumptions underlying the programme developers’ intervention plan are confirmed or refuted by the cluster’s actual development. Since the cluster development that occurs is practically always influenced by other factors as well, the influence of competing causal mechanisms on the outcome is equally subject to investigation. Given the key importance of the ‘programme theory’, impact evaluations that make use of PT are inevitably theory-led evaluations (Bamanyaki and Holvoet, 2016).

This raises the question which theory exactly should be considered. A closer look at German cluster programmes shows – and it is likely to be similar in other countries – that the programme makers are mostly vague about the underlying ideas on causal mechanisms that are to be set in motion by the planned intervention. Formulated programme targets are often quite general and diffuse, cannot be operationalised, and are sometimes even self-contradictory (Kiese, 2017; Rothgang and Lageman, 2021; Rothgang et al., 2019). However, this in no way means that the support is not based on a guiding principle. Similar to complex programmes in other policy areas (e.g. Birckmayer and Hirschon Weiss, 2000: 429), each cluster policy intervention is based on some notion of cluster-related theory, regardless of how profoundly or superficially this was considered by policy makers. Part of a PT analysis is to reveal this theory.

The following factors characterize the theoretical notions that underlie cluster programmes:

- Over the past decades, cluster research has developed a comprehensive body of knowledge on the economic benefits associated with clustering (e.g. Bathelt and Glückler, 2018). At the same time, innovation research has increased our understanding of the contribution of complex innovation networks to knowledge generation and diffusion (e.g. Powell and Giannella, 2011 [2010]). Today, research provides a very differentiated picture of the various manifestations of cluster policies, their remarkable resilience and the rationales underlying each of them (Uyarra and Ramlogan, 2017; Wilson, 2019). At the same time, there is no generally accepted theory of cluster policy.
interventions that could be based on a broad consensus of experts working in this field. In the absence of a generally accepted stringent cluster policy theory, we can at least identify a corpus of largely shared theoretical assumptions on the course, results and economic (dis-)advantages of industrial agglomeration processes. These assumptions are used by economists, geographers and sociologists, and constitute the elements of what can be denoted as a ‘cluster policy-related theory’ (descriptions, for example, in Andersson et al., 2004; Lindquist et al., 2013; Sölvell, 2009). That means, a basic consensus has developed among most cluster researchers on economic drivers of cluster development and the principal desirability of promoting cluster development under certain conditions.

- Although government programmes are (almost) always based on an explicit or implicit programme theory (Pawson, 2013: 87 f.), the relevant documents are the result of a negotiation process between the actors involved at the political and administrative levels who are also influenced by their personal and institutional biases. They have different ideas, which often change over time, regarding the impact mechanisms and effectiveness of the funding instruments to be used. In contrast to the corresponding ideas of specialist academics, pragmatic criteria such as robustness, simplicity and chance of wider public acceptance are paramount when writing a draft programme, while theoretical coherence is of secondary importance.

- In view of the general popularity of cluster and network policy approaches among political decision-makers in Europe and North America, one could speak of the silent presence of a folk cluster theory in public discourse. Of course, this is mainly induced by the absence of explicit statements in programme documents regarding the rationale for interventions. Political decision-makers who develop cluster programmes at least share a widespread idea about the usefulness of cluster policy, even though a deeper engagement of practitioners with relevant academic work hardly ever takes place.

Under these circumstances, a practical task for evaluators is to construct a theoretical narrative that exhibits the intentions of the programme. They need to explicitly formulate the theoretical assumptions about the causal mechanisms which lead to the desired impacts of the cluster programme. This theoretical guiding idea should do justice to the complexity of the programme effects and, at the same time, be as simple as possible. It should concentrate on essential points and be presentable in a logical framework that describes the programmes’ logic, stimuli, activities, chain of effects (intermediate results) and desired final impacts (American Evaluation Association 2015: 18 ff.).

After this theoretical framework is established, PT can be used to investigate whether the theoretically assumed causal mechanism was effective. At the same time, the evaluators can determine whether the intervention produced any desirable or non-desirable unintended results or whether the cluster development identified was caused by completely different mechanisms.

**Possible objects of investigation of PT in cluster policy impact evaluations**

The application of PT in impact evaluations can take two forms: On one hand, PT can serve as an overarching framework for impact evaluation into which various qualitative and quantitative components of investigation can be integrated. On the other hand, PT can be used as one
component of a more complex study design. Thus, the development of a sub-area of the general study object can be analysed, starting from a baseline situation and extending over several development stages up to the (previously defined) final state.

Consequently, PT can refer not only to the development of the cluster as a whole but also to the evolution of any part, single unit or sub-unit of it. We will demonstrate this with a few examples, each of which represents a different level of observation of cluster analysis.

Level 1: Development of a sponsored technology *project family*
Level 2: Development of the cluster *innovation network*
Level 3: Accompanying tracking or reconstruction of the development of the *cluster initiative* (organization)
Level 4: Development of a supported *cluster* starting from the baseline situation

For these four levels, general process characteristics can be derived. These characteristics are helpful in the application of PT at different levels. Table 1 displays the possible focus of the PT, process characteristics on the different levels and categories that can be used to assess the results in respect to different dimensions of impact.

| Level  | Focus of analysis | Process characteristics | Results |
|--------|-------------------|-------------------------|---------|
| 1:     | Project families  | single project as part of larger, interdependent project families; evolution of knowledge development; uncertainty of project success (technical, economic) | new technological knowledge, patents, additional revenue, employment |
| 2:     | Innovation network | often temporary network structure (determined by technological and market-related considerations of the actors); influenced by government policies; however, often only medium to short term | medium- to long-term development of strategic research and innovation networks (qualitative, quantitative); influence of government policies on networks; patterns of knowledge development, exchange |
| 3:     | Cluster initiative | development of cluster initiative; common activities; cluster management organization | initiated projects; impulses to individual actors; initiating common activities; knowledge exchange |
| 4:     | Cluster           | factors influencing cluster development – development of single firms, universities and research organizations, and their cooperation patterns; individual decisions by actors; framework conditions in respect to market development, technology | development of employment, turnover; initiated projects; impulses to individual actors; initiating common activities; knowledge exchange |
In the project level (level 1), individual projects are usually part of project families. Project ideas often develop in other project contexts and give impulse to new projects. While the individual project has clear goals, knowledge development takes place in the background of R&D activities. Publicly funded projects often involve projects across different actors (firms, research organizations, university departments). PT can follow the paths of related R&D activities to identify how projects funded by a cluster policy intervention influence the path of knowledge development, lead to follow-up projects, and finally to new products and changes in production processes or the establishment of start-up firms.

On a more aggregate level, the actors within a cluster are related to each other in innovation and research networks (level 2). These innovation networks are partly regional, but relevant network relations are also national or international. Publicly funded cluster programmes like other innovation programmes often require cooperative projects and, thus, influence the network structure. This influence can not only be short or medium term but also long term. The networks are important for the use of newly developed projects and for learning at the level of individual actors (firms, research organizations, universities). Several factors like different organizational cultures, different time structures in organization-internal processes and regional characteristics like industry structure can pose barriers to network development and knowledge exchange.

As intermediary organizations, cluster organizations (level 3) are characterized by rather specific traits of organizational development. Important aspects are the (degree of) involvement not only of single organizations but also of individuals on decision boards, means of public funding and activities of the cluster management as part of the cluster organization. While the formal organization is often not crucial for the success of cluster organizations, the attractiveness for individual firms/research organizations/universities to cooperate is a crucial factor. Public funding and the time pattern that relates to public programmes is crucial to understand how cluster organizations evolve (Rothgang and Lageman, 2017). Even though cluster organizations involve many actors, the goals and strategic developments are usually determined by only a few actors. Changes in the composition of these core actors can potentially influence the development of a cluster initiative. Changes can especially occur through strategic decisions of individual firms. Often, the methods used in impact evaluations are not suited to highlight the role of cluster organizations for the cluster’s development and the impact of the programme. PT is especially suited to identify how different factors (public funding, cluster management decisions, individual actors or persons) influence the development of a cluster organization over time and, thus, also its influence on the cluster.

The scope of the cluster (level 4) is broader than the cluster organization. It comprises all actors that cooperate in an industry or technology field within a region. While the core goal of cluster policies is to influence the development of individual clusters, it reaches only a fraction of the cluster directly. Thus, cluster development also is influenced by multiple other factors, among them international competition in the respective industry or technology field, the institutional/legal framework as well as the competitiveness and strategic decision of actors within the cluster. The PT approach can analyse the interdependence of these factors over time. It is also suited to uncover influence at other levels of analysis of cluster development.
The role of PT as component of evaluation designs

The key feature of PT is that it allows evaluators to analyse the causal mechanism assumed in the cluster policy programme as well as to trace the influences of alternative causal mechanisms. PT in evaluation contexts can be used for triangulating the results of quantitative and qualitative methods, which are linked by the interest in a common object of investigation. This is an essential difference to most ‘classical’ applications of PT in its disciplines of origin, where it is primarily a qualitative method of investigation. Recently, however, major authors have been open to combining PT with quantitative methods within the framework of mixed-methods designs (e.g. Checkel and Bennett, 2015: 272 f.).

In the case of the evaluation of a cluster promotion programme, PT should be able to trace the development of a cluster, considering all significant events that occurred during the period under study. The starting points of reference are the causal theory underlying the intervention, the programme makers’ expectations with regard to the development of the clusters and the corresponding indicators used to measure success. At the same time, other significant influencing factors have to be taken into account that are not connected with the intervention, but causally influence the cluster’s development. In combination with the results of PT, all results of the other research methods used in the evaluation, for example, randomized control trials (RCTs), can contribute to an understanding of causality. Based on regression analysis, RCTs provide findings on building blocks of cluster development, such as the influence of the intervention on the research expenditures of the clusters’ actors involved, findings on the development of research expenditures, patent applications, company sales, exports of cluster firms and so on.

A major advantage of the application of PT to cluster impact evaluation is that the strengths of individual case studies can be combined with those of the direct comparison of the programme-induced development of different promoted clusters (e.g. Beach and Rohlfig, 2018). Cluster programmes in many cases target several promoted clusters that share a programme specific promotion criterion: they may, for instance, either represent an emerging technology which is considered by experts to have substantial future potential, belong to an industrial sector considered eligible for support or form a spatial focal point of new industries in an old industrialized region. Even in the promoted clusters of the same industry or technology field, key contextual factors will almost always be very different. Important contextual factors are, for instance, the role of large and small enterprises and research organizations, the structure and growth potential of the regional economy, or the strengths and weaknesses inherent in the procedural mechanisms implemented by the cluster initiative.

In summary, the full potential of each method can only be realized by combining them into a meaningful whole based on the integration of individual results by triangulation (an example is Bjurulf el al., 2012). The shared object of investigation determines the synopsis of the results of different investigation methods and their causal interpretation. The challenge in evaluating complex systems is that observed states practically always have a multi-causal explanation. It is hardly conceivable that the development of a cluster can be traced back to the one original intervention that is evaluated. PT can play an important role in placing an isolated programme stimulus in the context of a variety of relevant influences.
**Summary and discussion**

The purpose of this article is to draw attention to an evaluation approach that could close a gap in the methodological repertoire for evaluating complex cluster programmes and that has so far only been applied in a few individual cases. The PT approach does not make any of the methods which are currently used in impact evaluations obsolete, but rather complements them. In this sense, PT should be understood as an additional component of research designs of evaluation studies, suitable to bring hitherto neglected aspects of the evaluation task into focus. It could help to direct the evaluators’ attention to the complexity of the object being promoted, avoiding the common mistake of establishing the impact of a programme on the selective findings obtained by applying a single preferred analytical tool. A particular strength of this research approach is its flexibility in adapting to different research tasks such as prosessual analyses of different levels of cluster development (e.g. project families, innovation networks, cluster initiatives, clusters).

Interventions in cluster development are highly dependent on the conditions of their respective sectoral and regional environments as well as of the local innovation systems in which cluster actors operate (Bramwell et al., 2019). PT is particularly useful to illustrate the regional, local and situational contexts of cluster programme interventions. The application of PT in the evaluation of larger cluster programmes opens the possibility of comparing the chains of effects and impacts of the programme in different industry and regional environments as well as other idiosyncratic preconditions (including the role of individual core actors active in a cluster initiative).

PT can make a contribution to the understanding of the systemic character of cluster policy interventions. However, it is not a means to capture the systemic dimension of cluster development. The systemic challenge of evaluating complex programmes remains largely unresolved, at least in the case of cluster programmes. A solution would require suitable system-oriented evaluation methods to capture the system dynamics, which neither trivialize the systemic interrelationships nor get lost in increasingly complex model constructions that are useless for practical evaluation purposes.

Much of what has been said here may sound familiar to those scholars who are familiar with the evaluation activities of government programmes in Anglo-Saxon countries. This applies above all to those areas of government action that are not primarily concerned with economic analyses of the impact of structural policies such as, for instance, the health policy in the United Kingdom (Medical Research Council (MRC), 2019). However, such considerations have so far not been widely acknowledged in economic impact analyses of cluster policy. One important reason for this may lie in the idiosyncratic methodological thinking and practical approaches of researchers of different disciplines who are engaged in evaluation research.

An important argument in this article is that PT is a complementary research approach well suited for the analysis of some key aspects of the impact of cluster policy interventions which so far have not received the attention they deserve. There are also limitations to our discussion that need to be addressed in the future. One is that the proposal to include the approach in complex evaluation procedures had to be very general. Another limitation is that methodological and epistemological background aspects – such as suitable procedures for quality assurance of the causal analyses or the philosophical interpretation of causal mechanisms – could only be touched on in passing. Also, in the given context, it was not possible to illustrate the application of PT with specific examples of cluster policy impact evaluations.
Acknowledgements
We thank Martina Böhmel and Ida Zinke for their assistance in literature search and completion of the manuscript.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iD
Michael Rothgang https://orcid.org/0000-0001-8328-0295

Note
1. ‘Replication crisis’. Retrieved from: https://en.wikipedia.org/w/index.php?title=Replication_crisis&oldid=983969177, accessed 17 October 2020.

References
Abbott A (2001) Time Matters. On Theory and Method. Chicago and London: The University of Chicago Press.
American Evaluation Association (2015) Evaluating outcomes of publicly-funded research, technology and development programs: Recommendations for improving current practice (Version 1.0). Report, February. Research, Technology and Development Evaluation Topical Interest Group of the American Evaluation Association (AEA). Available at: www.eval.org (accessed 21 January 2020).
Andersson T, Schwaag S, Sörvik J, et al. (2004) The Cluster Policies Whitebook. Holmberg: IKED International Organization for Knowledge Economy and Enterprise Development.
Bamanyaki PA and Holvoet N (2016) Integrating theory-based evaluation and process tracing in the evaluation of civil society gender budget initiatives. Evaluation 22(1): 72–90.
Barnett C and Munslow T (2014) Process tracing: The potential and pitfalls for impact evaluation in international development – Summary of a workshop held on 7 May 2014. Evidence report no. 102. Brighton: IDS.
Bathelt H and Glückler J (2018) Wirtschaftsgeographie: Ökonomische Beziehungen in räumlicher Perspektive. 4., vollst. überarb. u. erw. Auflage. Stuttgart: Eugen Ulmer.
Beach D and Rohlfing I (2018) Integrating cross-case analyses and process tracing in set-theoretic research: Strategies and parameters of debate. Sociological Methods and Research 47(1): 3–36.
Beach R and Pedersen RB (2019) Process-tracing Methods. Foundations and Guidelines, 2nd edn. Ann Arbor, MI: The University of Michigan Press.
Befani B and Stedman-Bryce G (2017) Process tracing and Bayesian updating for impact evaluation. Evaluation 23(1): 42–60.
Bennett A and Checkel JT (eds) (2015a) Process Tracing: From Metaphor to Analytic Tool: Strategies for Social Inquiry. Cambridge: Cambridge University Press.
Bennett A and Checkel JT (2015b) Process tracing: From philosophical roots to best practices. In: Bennett A and Checkel JT (eds) Process Tracing: From Metaphor to Analytic Tool. Cambridge: Cambridge University Press, 3–37.
Bergman EM (2008) Cluster life-cycles: An emerging perspective. In: Karlsson C (ed.) Handbook of Research on Cluster Theory. Cheltenham; Northampton, MA: Edward Elgar, 114–32.

Birckmayer JD and Hirschon Weiss C (2000) Theory-based in evaluation in practice. Evaluation Review 24(4): 407–31.

Bjurulf S, Vedung E and Larsson CG (2012) A triangulation approach to impact evaluation. Evaluation 19(1): 56–73.

Brady H and Collier D (eds) (2004) Rethinking Social Inquiry: Diverse Tools, Shared Standards. 1st edition. Rowman and Littlefield.

Bramwell A, Hepburn N and Wolfe DA (2019) Growing entrepreneurial ecosystems: Public intermediaries, policy learning, and regional innovation. Journal of Entrepreneurship and Public Policy 8(2): 272–92.

Centre of Excellence for Evaluation Government of Canada (CEE) (2012) Theory-based Approaches to Evaluation: Concepts and Practices. Ottawa, ON, Canada: CEE. Available at: https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/evaluation-government-canada/theory-based-approaches-evaluation-concepts-practices.html

Checkel JT and Bennett A (2015) Beyond metaphors: Standards, theory, and the ‘where next’ for process tracing. In: Bennett A and Checkel JT (eds) Process Tracing: From Metaphor to Analytic Tool: Strategies for Social Inquiry. Cambridge: Cambridge University Press, 260–75.

Collier D (2011) Understanding process tracing. PS – Political Science and Politics 44(4): 823–30.

Coryn CLS, Noakes LA, Westine CD, et al. (2011) A systematic review of theory-driven evaluation practice from 1990 to 2009. American Journal of Evaluation 32(2): 199–226.

DeGEval (2017) Standards für Evaluation. Erste Revision 2016. Mainz: DeGEval – Gesellschaft für Evaluation.

Dohse D (2000) Technology policy and the regions: The case of the BioRegio contest. Research Policy 29: 1111–1133.

Edler J and Fagerberg J (2017) Innovation policy: What, why, and how. Oxford Review of Economic Policy 33(1): 2–23. https://doi.org/10.1093/oxrep/grx001

Fornahl D, Heimer T, Campen A, et al. (2015) Cluster als Paradigma der Innovationspolitik: Eine erfolgsreiche Anwendung von Theorie und Praxis? Studien zum deutschen Innovationssystem 13-2015. Bremen; Frankfurt am Main: CRIE – Centre for Regional and Innovation Economics, Universität Bremen, Technopolis Deutschland GmbH.

Fromhold-Eisebith M and Eisebith G (2008) Looking behind facades. Evaluating effects of (automotive) cluster promotion. Regional Studies 42(10): 1698–719.

Gerring J (2008) The mechanismic worldview: Thinking inside the box. British Journal of Political Science 38(1): 161–79.

Gerring J (2019) Case Study Research. Principles and Practices. Strategies for Social Inquiry. Cambridge: Cambridge University Press.

Gertler PJ, Martinez S, Premand P, et al. (2016) Impact Evaluation in Practice. 2nd Edition. Washington DC: The World Bank. Available at: https://worldbank.org/ieinpractice (accessed 22 August 2021).

Giel S (2013) Theoriebasierte Evaluation. Konzepte und methodische Untersuchungen. Internationale Hochschulschriften 584. Münster: Waxmann.

Grant A, Treweek S, Dreischulte T, et al. (2013) Process evaluations for cluster-randomised trials of complex interventions: A proposed framework for design and reporting. Biomed Central. Trials 2013: 1415.

Hedström P and Ylikoski P (2010) Causal mechanisms in the social sciences. Annual Review of Sociology 36: 49–67.

Karlsson C (ed.) (2008) Handbook of Research on Innovation and Clusters: Cases and Policies. Cheltenham; Northampton, MA: Edward Elgar.

Kiese M (2017) Regional cluster policies in Germany: Challenges, impacts and evaluation practices. Journal of Technology Transfer 44: 1698–719.
Krugman P (1997) Development, Geography, and Economic Theory (ed Pbck). Cambridge, MA; London: The MIT Press.

Lehmann EE and Menter M (2017) Public cluster policy and performance. Journal of Technology Transfer 43: 558–92.

Lindquist G, Ketels C and Sölvell Ö (2013) The Cluster Initiative Greenbook 2.0. Stockholm: Ivory Tower Publishers.

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

Medical Research Council (MRC) (2019) Developing and evaluating complex interventions: Following considerable development in the field since 2006, MRC and NIHR have jointly commissioned an update of this guidance to be published in 2019. Available at: https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/

Pawson R (2013) The Science of Evaluation: A Realist Manifesto. Los Angeles, CA: SAGE.

Porter M (1998) Clusters and the new economics of competition. Harvard Business Review. Available at: https://hbr.org/1998/11/clusters-and-the-new-economics-of-competition

Powell W and Giannella W (2011 [2010]) Collective invention and inventor networks. In: Hall BH and Rosenberg N (eds) Handbook of the Economics of Innovation, vol. 1. Amsterdam: Elsevier, 575–605.

Rothgang M and Lageman B (2017) Do cluster initiatives develop in cycles? The example of Germany’s spitzenclusters. In: Fornahl D and Hassink R (eds) The Life Cycle of Clusters: A Policy Perspective. Cheltenham: Edward Elgar, 222–41.

Rothgang M and Lageman B (2021) Why are there so few hard facts about the impact of cluster policies in Germany? A critical review of evaluation studies. Review of Evolutionary Political Economy 2: 105–29.

Rothgang M, Dehio J and Lageman B (2019) Analysing the effects of cluster policy: What can we learn from the German leading-edge cluster competition? The Journal of Technology Transfer 44: 1673–97.

Sölvell Ö (2009) Clusters. Balancing Evolutionary and Constructive Forces. Stockholm: Ivory Tower Publishers.

Uyarra E and Ramlogan R (2016) The impact of cluster policy on innovation. In: Edler J, Cunningham P, Gök A and Shapira P (eds) Handbook of Innovation Policy Impact. London: Edward Elgar, 197–231.

Uyarra E and Ramlogan R (2017) Cluster policy in an evolutionary world? Rationales, instruments and policy learning. In: Fornahl D and Hassink R (eds) The Life Cycle of Clusters: A Policy Perspective. Cheltenham; Northampton, MA: Edward Elgar, pp. 35–55.

Wilson JR (2019) Cluster policy resilience: New challenges for a mature policy. International Journal of Business Environment 10(4): 371–82.

Michael Rothgang is Researcher at the RWI-Institute for Economic Research. He coordinated several large technology evaluation projects. His research fields are cluster development and policy, technology transfer, and technology policy evaluations.

Bernhard Lageman was Head of the research department ‘Firms and Innovation’ at the RWI-Institute for Economic Research until his retirement in 2013. Today, he is Associate Researcher at the RWI. His major research interests are science and technology studies.