fteval JOURNAL
for Research and
Technology Policy Evaluation

APRIL 2022, Vol. 53, pp. 181-191
DOI: 10.22163/fteval.2022.556

SYSTEMS ANALYSIS IN INNOVATION POLICY EVALUATION THE UNFULFILLED PROMISE

MICHAEL ROTHGANG AND BERNHARD LAGEMAN
DOI: 10.22163/fteval.2022.556

ABSTRACT

Our paper addresses the question, why systemic approaches have only played a modest role in impact evaluations of innovation and technology programmes so far and examines possible reasons for this shortcoming, as well as discussing solutions that could be offered to remedy the existing deficit. While the need for a systemic approach to evaluations has been stressed quite often, the methodological challenges and reasons for the lack of systemic evaluations in practice have to our knowledge not yet been addressed in a systematical manner. This contribution is conceptual in nature and based on a review of the research literature on the use of systemic approaches in evaluations of the impact of R&I policy programmes. The analysis shows that the use of systemic methods encounters both epistemological and institutional obstacles. Suggestions are made for the further development of the methodological repertoire by including suitable systemic approaches.

1 INTRODUCTION

Systems analysis, a child of the development of applied sciences around the Second World War, has experienced periods of boom and bust over the past seven decades (Barbrook-Johnson et al. 2021; Williams 2015). Today, in view of climate change and the development of global IT networks and artificial intelligence, the prevailing circumstances seem more favourable than ever for a broad acceptance of systems thinking and complexity research. The increased attention that is given to systems thinking[1] is not only the result of the insightful model-based analyses of environmental studies (Meadows 2008), but also serves as a vehicle for spreading awareness of globally coordinated efforts to curb atmospheric warming (e.g., Ison and Straw 2020).

Innovation systems research, which was particularly driven by Scandinavian authors (e.g., Chaminade et al. 2018; Edquist 2005; Lundvall 2010), has made a significant contribution to the development of innovation economics in the closing decades of the 20th century. It emerged mainly against the background of the insight into the fundamental shortcomings of the "linear model of the innovation process", which characterised the views of outstanding theorists in the 1940s and 1950s when innovation economics was still in its formative phase (e.g., Bush 1960 [1945]). The system concept, originally geared to the study of national innovation systems, was successively narrowed down to more limited areas of the economic sphere which form subsystems of the broader national innovation system: sectoral and regional innovation systems, technology fields, and industrial clusters. Moreover, in the context of the recent "transformative turn" in the innovation policy literature, systems approaches have gained a strong momentum (e.g., Borrás and Edler 2020; Schot and Steinmueller 2018).

Against this background, it is surprising that systems thinking has never taken firm root in the policy areas responsible for technology and innovation, i.e., has not been properly internalised by decision-makers. As Borrás and Edquist (2019: 40-42) observe, most actions of innovation policy in industrialised countries are based on the linear model. They offer the explanation that the linear model is capable of conveying the impression of a clear, easily comprehensible causal connection between policy measures and their intended effects. While the theoretical research on national innovation systems has found great resonance in innovation research, its practical consequences seem rather difficult to grasp and manage in the political process or to communicate to the public.

This paper discusses the extent to which systemic thinking has penetrated the practice of evaluating innovation policy programmes to date, and explores causes of what we perceive to be a relatively weak response in evaluation practice to systems and complexity theory and innovation systems research to date. We focus on the application of systems thinking to the evaluation of individual innovation programmes that fulfil elementary complexity criteria. In contrast, Borrás/Laatsit (2019) analyse diffusion of system thinking and systemic evaluation practices in the innovation policy field of the EU28 in total. The following discussion is strongly influenced by the experience gained by the authors in Germany. However, we assume that comparable experiences can also be confirmed by a closer analysis of the innovation policy-oriented evaluation practice of other continental European countries. We argue that

1 Systems analysis is inevitably linked to research into the system complexity of the systems under study - the central property of complex adaptive systems. This should be taken into account when we refer to "systems thinking" and "systems analysis" in the following. In the following, we focus on systems, because studies that claim to capture the complexity of the object of study in one way or another do not necessarily have to adopt a systems perspective on the object, i.e., they do not have to focus on its dynamic whole.

due to the flexibility and breadth of the relevant methodological tools, unused potentials of system-oriented evaluation research could be activated without critical additional expenditure of human and financial resources.

The paper proceeds as follows: In section 2, we take a look at the results of a Scopus query of the coverage of scientific papers on systems analysis and systems evaluation (section 2). In section 3, we address the diffusion process of systems thinking in evaluation research in general as well as in innovation policy and its evaluation. Subsequently, complex innovation programmes are introduced as an object of investigation (section 4) and complexity attributes are demonstrated at the example of two German programmes (section 5). In section 6, we ask, what a systems evaluation is and address possible reasons for the low degree of reception of systems thinking in the evaluations in section 7. Section 8 concludes with practical suggestions for a pragmatic handling of systems approaches.

2 SYSTEMS ANALYSIS AND SYSTEMS EVALUATION PAPERS IN SCOPUS

The systems concept received considerable attention in innovation research, but it never became the dominant paradigm. The ground for a more systemically oriented evaluation practice in the field of innovation policy was therefore rather rocky in the beginning. In this context, the first question is to what extent systems thinking has influenced evaluation theory and evaluation research practice in general.

If systems thinking were to be used to a greater extent in evaluations of innovation policy programmes, this should have made itself felt in the relevant contributions to scientific journals. However, it should be noted that evaluations, which in most cases are commissioned research, do not necessarily result in journal articles. Nevertheless, a considerable and, it can be assumed, also representative part of practical evaluation research sooner or later finds its way into the world of scientific journals.

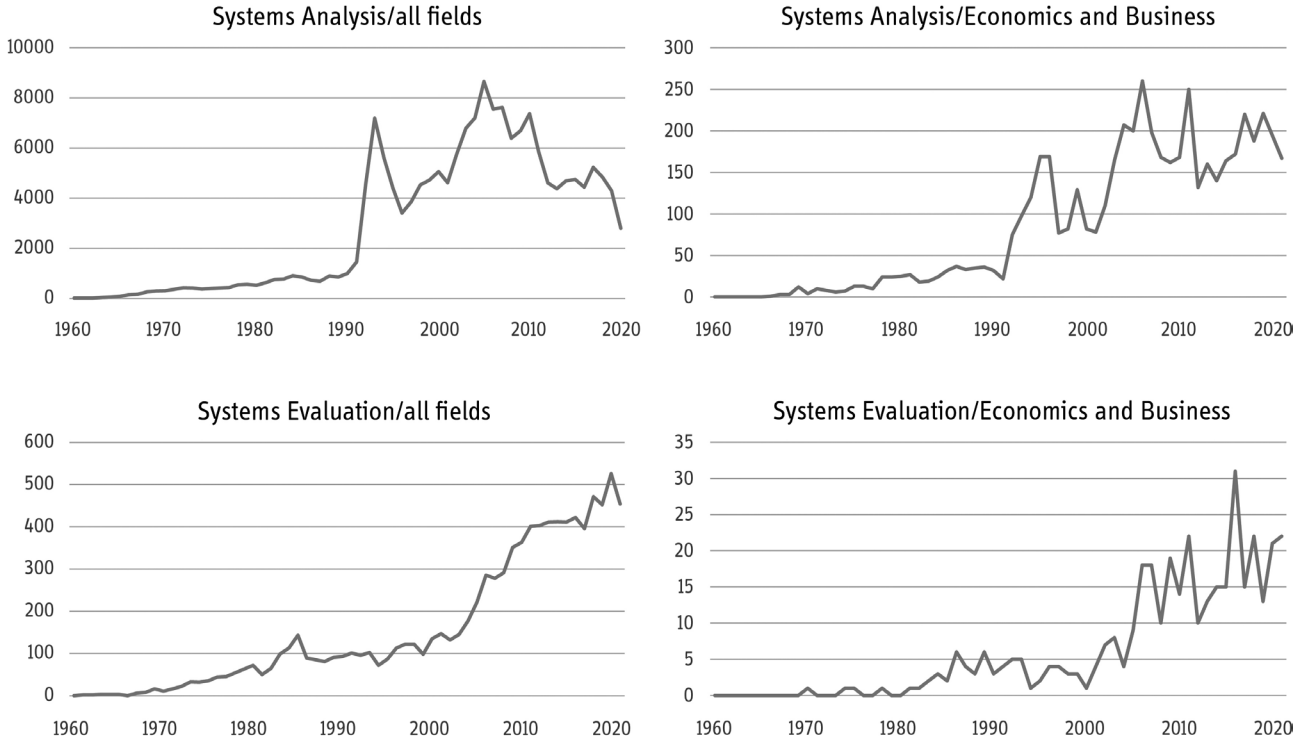
A recent search of the scientific bibliographic database Scopus showed that systems analysis, beginning in the 1960s, has found its way into the scientific literature. A strong upswing has been observed since the early 1990s, which reached a first peak in 1994 with 7.197 registered titles and its highest level to date in 2006 with 8.635 mentions (Figure 1, upper left side).

If one narrows down the search to the fields of "economics" and "business", comparatively low numbers emerge. For the period from 1960 to 2021, the highest value was reached in 2006 with 260 titles (Figure 1, upper right side). This corresponded to a share of 3.0% of all titles listed in connection with systems analysis. The highest share value was registered in 1978 with 5.6% (24 titles). A look at the predominant keywords of the registered titles shows that they are predominantly assigned to the fields of business administration/management, operations research and IT development.

In addition, the results of a Scopus search for titles containing the keywords "system(s) evaluation" or "systemic evaluation" are given in Figure 1 (lower side). The number of titles has increased significantly since the turn of the millennium. An analysis of the keywords shows a

Figure 1: Systems analysis and systems evaluation in the research literature as recorded in the database Scopus, all disciplines and economics/business, number of registered publications, 1960-2021

Source: Own data search and depiction; keywords "system analysis"/"systems analysis"; "system evaluation"/"systemic evaluation" in all areas and in the areas "economics" and "business" in the fields "title", "abstract", "keywords", 26 October 2021.



clear dominance of works that can be assigned to computer science in the broadest sense. As Figure 1 (lower right side) shows, the number of titles from the fields of "economics" and "business" is rather small, but it also increased substantially in the past decades. It reached its peak in 2016 with 31 registered works (7.4 % of all publications on “system(s) evaluation”).

Our attempt to identify articles with the keywords "systemic evaluation" or "systems analysis" and "innovation policy" ended up with a short list of 9 titles in total for the entire period under review. The central result is, in other words, that there is no evidence for a more widespread discussion of systems evaluation in scientific papers in the field of innovation policy. We assess the relevance of this observation by taking a look at the general uptake of systems thinking on the one hand in policy evaluations in general and on the other hand in the field of innovation policy evaluations in the following section.

3 SYSTEMS THINKING IN (INNOVATION) POLICY EVALUATION

Theoretically oriented evaluation research has taken up the impulses of systems theory and its sibling complexity theory since at least the early 2000s. In the meantime, both streams of theory development have also found their way into practical evaluation work (Midgeley 2006). Admittedly, this applies more to the Anglo-Saxon countries than to continental Europe. The pioneering work of the Santa Fe Institute, New Mexico, founded in 1984 and the recent evaluation-oriented research of the UK Centre for the Evaluation of Complexity across the Nexus (CECAN) as well as the inclusion of complexity-related evaluation approaches into the 2020 edition of the Magenta Book (HM Treasury 2020, for the work of CECAN see the special issue of Evaluation, 1/2021, dedicated to the topic of complexity) are outstanding examples of the increased attention to systems approaches in evaluation research. Another milestone in the dissemination of systems thinking in evaluation research is the “Expert Anthology”, published in 2006 by the American Evaluation Association (Williams and Imam 2006).

Evaluation research, as a branch of applied knowledge that is constantly striving to absorb new concepts, should actually have played a role as a pioneer of systems thinking in politics. Indeed, such a reception of systems theory thinking can be observed in evaluation research, albeit rather late. When the CECAN authors talk about a “turn towards complexity” (Barbrook-Johnson et al. 2021: 5; earlier use of the term in Mowles 2014), this applies not only to theoretical research on policy evaluation tasks, but also to parts of the practical evaluation activities in the Anglo-Saxon countries. However, even here, policy evaluations are much more focused on health care and CECAN's “nexus”, namely food production, energy production, water management and the handling of environmental issues, than on the area of interest to us here, innovation policy. In contrast, a recent policy paper of the American Evaluation Association's Research, Technology and Development Evaluation Group (AEA 2015), to cite a current example, does not specifically address the possible use of a systemic perspective in the evaluation of R&D programmes.

In particular, those policy areas in which the problem pressure resulting from the perception of the complexity of the tasks to be solved was particularly substantial proved to be receptive to systems thinking. Prominent examples are environmental, health and development policy. The evaluation field in which the system perspective has probably gained the strongest foothold so far is the evaluation of development projects (e.g., Williams 2015 seems especially inspired by this field). This can be explained by the fact that in no other policy field is the pressure for a close examination of the impact of projects as great as in the field of development aid. After all, there are "lost decades" of many failed development aid projects in the poorest countries of the world, in which not only misconceived large-scale projects, but also many well-conceived manageable endeavours have proven to be failures in retrospect (Easterly 2007; Moyo 2009). The increasing insistence on conducting rigorous evaluations, preferably experimental designs and Randomized Control Trials (RCTs), where they are possible, as well as the strong emphasis on introducing a systemic perspective into the evaluation of projects, can be seen as a reaction to earlier failures.

But how is the present situation in the field of innovation policy evaluation? Unfortunately, there is no database on evaluation studies that could hold a candle to Scopus. The evaluation study database SIPER could possibly provide valuable information on the extent to which systemic thinking has found its way into the practical evaluation of innovation policy measures. Unfortunately, the search criteria do not yet allow external users to search for relevant keywords. A cursory review of evaluation studies recorded in this database shows at least that many evaluation designs contain methodological components that could also be part of system perspective evaluations. However, as will be shown below, this alone does not constitute a system evaluation.

This also applies to evaluations in the national framework of Germany, where the authors have practical experience in the evaluation business: We are not aware of any evaluation of a relevant programme that has explicitly committed to a systemic evaluation approach, although this would of course always be theoretically possible within the framework of a mixed method design.

At the same time, however, there is a certain unease among researchers. It can by no means be said that evaluation researchers are oblivious to the actual complexity of their objects of investigation when working on evaluation assignments. Researchers who have been in the field of evaluating complex state interventions are familiar with the gut feelings that structural and process characteristics of the impact patterns of innovation programmes are actually much more complicated than they appear in the indicators and methods used. Practitioners can only point out in their studies that there are still many relevant influencing factors and impact mechanisms that elude analytical access for the time being.

Innovation researchers have also recently repeatedly called for systemic evaluations of innovation policy. Edler/Fagerberg (2017) emphasise that “the available evidence on innovation policy impacts at the national level seems to suggest that holistic – or systemic perspectives in policy is important” (p. 14). The authors have primarily the call for cross-instrumental evaluations in mind when they emphasise that the overwhelming majority of evaluations to date have focused on a single instrument (2017: 13). In their comprehensive study of evaluation practices in EU28 Borrás/Laatsvit (2019) found that few countries have developed a system-oriented type of innovation policy evaluation. However, the lack of systemic analysis also applies to individual programmes that address a complex object of investigation such as clusters. As Uyarra/Ramlogan (2017) point out, the research methods used today are hardly capable of

adequately capturing the complex interactions of a multitude of actors who dominate this policy field.

4 COMPLEX INNOVATION PROGRAMMES AS OBJECT OF INVESTIGATION

It would hardly be possible to provide a binding uniform definition of the concept of system in view of the ubiquitous presence of dynamic systems in all areas of reality, and such a definition does not exist (so for example also Williams and Hummelbrunner 2011: 16). The meticulous, comprehensive attempt made by Ackoff (1971) half a century ago to define the system concept as precisely as possible is hardly suitable for capturing the diversity of what is rubricated under "systems research" today.

Essential elements of a pragmatic definition for practical use, however, that can be found throughout the contributions of the various authors are: (i.e., Ackoff 1971; Mainzer 2008, 2015; Meadows 2008): A system is a set of interconnected elements from any realm of physical or virtual reality that form a whole, which are in mutual dynamic relationships with each other and in their interaction can produce properties of the whole that are not inherent to the individual elements. This whole - the system - exhibits certain changing patterns of behaviour (roles or functions) that determine its interactions with its constantly changing environment.

The central property of the type of system we are interested in here – complex adaptive systems - is their increasing complexity over time, which guarantees their survival in a dynamic world requiring their balance between order and chaos (on complexity cf. Holland 2014; Mainzer 2008; Mitchell 2009). Therefore, the topic we are interested in here is analysed on an abstract basis both in systems theory and the related field of complexity theory. Complex adaptive systems are characterised by (e.g., Forss and Schwartz 2917; HM Treasury 2020; Mainzer 2008, 2015; Meadows 2008):

- the continuous adaptation to challenges from the system's environment or internal relations,
- the occurrence of feedback loops in the development processes of the system,
- the appearance of non-linearities in the development of the system elements and the system,
- the ability of the emergence of new properties at higher levels within the system,
- the evolution of the system in a self-organized way,
- the existence of (non-deterministic) path-dependencies in the evolution of the system, and
- uncertainty and resulting limited calculability of the development of the system including the generation of unexpected effects.

The innovation system as a whole and components of it are complex adaptive systems. We denote all programmes that are intended to influence the evolution of the innovation system as a whole or of one of its components as "complex innovation programmes". These components can be, for example, sectoral innovation systems, technology fields, technology clusters, or innovation networks.

The complexity of the matter addressed is usually reflected in the complexity of the programme, for example in complex target bundles that may themselves contain trade-offs between individual targets. For the classification of a programme as "systemic-oriented" or "complex", it is not necessarily decisive whether the programme makers are fully aware of the complexity of the programme object.

5 SYSTEMIC CHARACTERISTICS OF TWO KINDS OF PROGRAMMES

A systemic perspective of an evaluation makes sense if the programme under investigation is directed at an area of intervention that has systemic qualities and the intervention thus addresses systemic goals. The criterion for introducing a systemic perspective cannot therefore be, as Imam et al. (2006) correctly state, whether a programme is financially voluminous or not, or whether the implementation process is complicated. In the following, we examine two examples from federal German innovation policy to demonstrate that systemic approaches can be implemented, on the one hand, in large programmes that are well equipped with financial resources and, on the other hand, in the case of small programmes that are provided with little funding. Both programmes mentioned here belong to a group of programmes that can be classified as systemic instruments "avant la lettre" according to the analysis by Smits/Kuhlmann (2004).

Table 1: Characteristics of two German innovation policy programmes at the federal level

	Industrial Collective Research (IGF)	"go-cluster"
Funding Ministry	Federal Ministry of Economic Affairs and Energy (BMWi)	Federal Ministry of Economic Affairs and Energy (BMWi)
Year/Date of establishment	1954	1 July 2012
Financial scope of the subsidy	169 Mio. € in 2018 for around 550 R&D projects; total sum approx. 4 bn € cumulatively since the year of establishment (2020 prices)	1 July 2012 – 30 June 2015: 1.5 Mill €; in 2018 626 vouchers for the improvement of innovation management in participating clusters; total sum spent in 2012-2020 amounts to approx. 4.75 Mill €
General promotion objective	Strengthening the research base of medium-sized industry	Providing a stimulus to improve cluster management of cluster initiatives that meet certain quality criteria to enable them to turn in highly effective international clusters
Mediator organisation(s)	AiF – German Federation of Industrial Research Associations & 101 Industrial Research Associations	VDI/VDE Innovation + Technik GmbH as project executing agency; Internet presence at "Clusterplattform Deutschland"
Grant recipients	Research organizations (institutes of research associations, university institutes, other institutes)	Participating cluster initiatives (at present 84)
Use of the funding for	Funding of industrial research and development projects carried out by independent research institutes or university institutes; promoted R&D projects should address the research needs of SMEs in particular	Advice and training for cluster managers; support of knowledge exchange between national and foreign cluster initiatives; support in the establishment of international contacts
Sectoral and technological orientation	Open to all technologies and industrial sectors; traditional focus more on highly developed conventional technologies, recently increased presence of high-tech sectors such as the IT sectors	Open to all technologies and industrial sectors; innovative industrial clusters are to be promoted above all, so the actual preference is more for new technologies

Source: Own depiction, data on "go-cluster" for 2012 – 2015 from Eckert et al. 2016: 76; data for 2018 from Deutscher Bundestag 2019: 4-5; for detailed information about the IGF programme cf. RWI and WSF 2010.

The more comprehensive of the two programmes is the programme for the promotion of industrial collective research (IGF), while the "go-cluster" programme is very modestly funded. Both address structural policy objectives and are intended to strengthen the competitiveness and innovative strength of the German economy. Table 1 provides an overview of the major parameters of the two programmes.

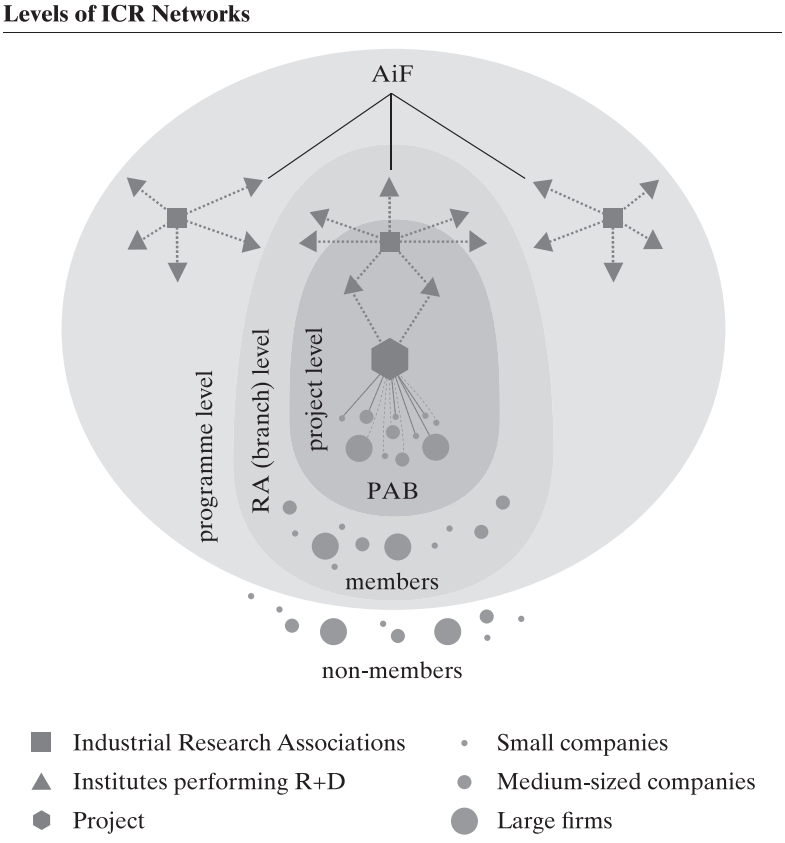
While the IGF focuses on funding collaborative research projects that serve specific research interests of companies and are often initiated by them in dialogue with research organisations (Figure 2), the go-cluster focuses on funding selected activities of the cluster management of "innovative clusters". In the case of IGF, the project proposals of the research associations are reviewed in a system-internal peer review process. In the case of go-cluster, the project management organisation reviews the applications of cluster initiatives that apply for participation based on an agreed catalogue of criteria. Successful participants are awarded one of the honorary labels (gold, silver, bronze) of the European Cluster Excellence Initiative (ECEI) in an internal selection process.

Both programmes aim to contribute to securing the long-term competitiveness of the German economy and thus, at least indirectly, to foster productivity, growth and job security. In both cases, the programme's aspirations go far beyond the immediate funding purpose and, in accordance with the programme logic, focus on central aspects of the long-term development of social welfare, although only in the case of the IGF these ambitions are backed up with substantial financial resources. In principle, both programme rationales can draw on the scientific authority of innovation economics research as well as cluster research. However, whether the programmes are actually suitable for producing the hoped-for positive effects on innovation and growth must - as always in such cases - be left to the results of evaluation studies.

What makes these programmes systemic? The answer to this question is to be found in the object of state intervention, in the actors on whose actions the success of the programmes ultimately depends, in the processes envisaged and the results aimed at:

- **Object of state intervention**
  According to the programme organisers (AiF), the IGF project funding is aimed at around 50,000 small and medium-sized enterprises and currently 101 research associations and research organisations are involved in industrial research. The IGF thus addresses an industrial research network that encompasses large parts of German industry. This comprehensive network consists of a multitude of nested individual networks (Figure 2) such as 101 industry level innovation networks[2] and hundreds of project level networks.
  Go-cluster (currently) supports 84 cluster initiatives that claim to represent a spatially located cluster consisting of vertically and horizontally connected companies, research organisations, and associated organisations.
- **Actors**
  In the case of both programmes, a large number of individual companies and research institutes are involved, whose development depends on a large number of individual decisions and internal and - only to a small extent influenceable - external influences. Thereby, the influence of the programme on the development of the individual firm in the vast majority of cases can only be of a marginal, hardly measurable dimension.
- **Processes**
  The development processes of the objects to be promoted are of an extremely diverse and complex nature and are in principle beyond the control of the state. They are characterised by feedback loops and are generally non-linear in nature. Emergence plays a role in the development of networks and clusters that are fostered by both programmes.
- **Results**
  The ultimate results of the complex state intervention, which occur over a long chain of indirect effects, are uncertain. They can neither be planned nor controlled in advance. In principle, considerable time elapses before the results of such an intervention can be realised. In the analytical identification of programme effects, an attribution problem has to be solved, as they are the result of the influence of multiple interdependent factors.

Figure 2: German Industrial Collective Research as system of systems



ICR – Industrial Collective Research (IGF)
AiF – Arbeitsgemeinschaft industrieller Forschungsvereinigungen (German Federation of Industrial Research Associations)
PAB – Project Advisory Board (Projektbeirat, members are firm representatives, accompanies individual research project)
RA – Research association (organizing pre-competitive research at sectoralindustry level)
Source: Rothgang et al. 2011.

2 The AiF is the umbrella organisation of 101 independent industrial research associations that organize pre-competitive research, each operating at sectoral industry level.

Thus, both programmes can be classified as systemic innovation policy instruments that can be adequately evaluated by employing systems evaluation approaches.

6 WHAT IS "SYSTEMS EVALUATION"?

We define "system evaluations" as evaluation approaches that give central importance to the investigation of the systemic interrelationships that determine the development of the object of study. They aim to do justice to the complexity of the object under investigation and are thus characterized by a holistic perspective. Central features of systemic evaluations are the critical reflection of the perspectives adopted in the analysis of the object and the definition of the boundaries of the system to be examined (this aspect particularly emphasised in Hummelbrunner 2011; Williams 2015). Systems evaluations are characterized by a perspective on the whole of government intervention concerned, which places the expected or observed effects of the intervention in the larger context of the respective policy and, on this basis, examines their meaningfulness and relevance.

Consequently, a systems evaluation is not a specific evaluation method that can be placed alongside other methods such as RCTs or peer interviews. It is not defined by the application of this or that exclusive method or, in the case of multi-method designs, of a specific set of methods. In principle, systemic evaluations can make use of the entire arsenal of quantitative and qualitative methods commonly used in evaluation research. Nevertheless, not every combination of methods is equally suitable to support systemic evaluation.

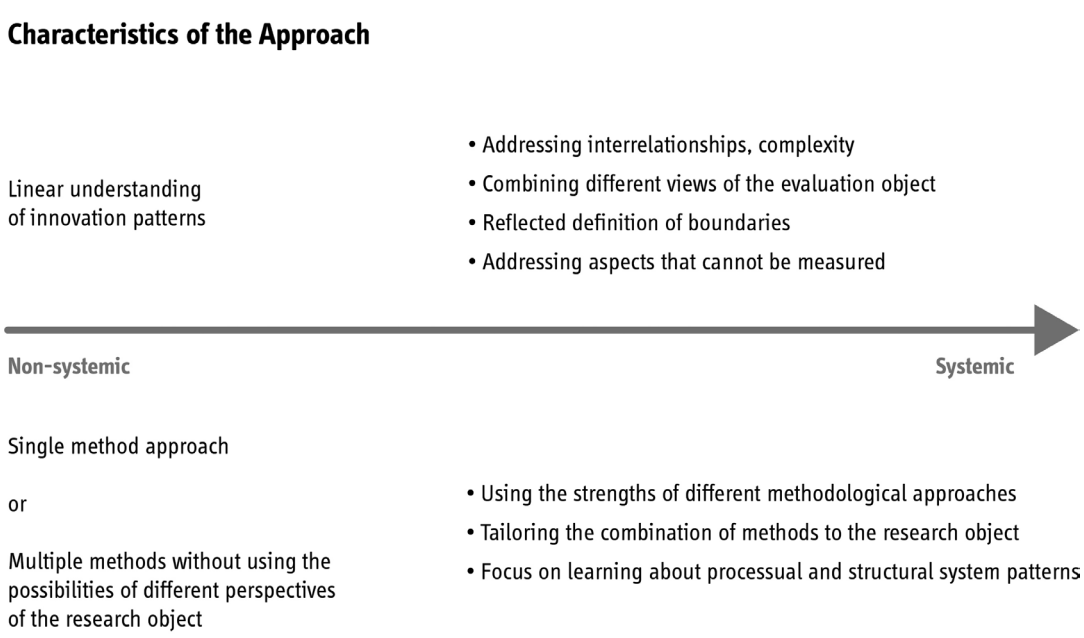
A systemic evaluation considers such development potentials that go beyond the framework of simple, linear causal relationships between the elements of the system. It is appropriate whenever the object to be evaluated is situated in a systemic context that is essential for understanding the effectiveness of the policy intervention and the mechanisms that it intentionally or unintentionally triggers.

The question arises, when an investigation takes on a "systemic character" or when does it definitely lose it? This question is relevant for all disciplines of scientific research that deal in any form with complexity. If the researcher opts for a rather generous answer, any investigation reveals systemic qualities. In order to bring a clear analytical line into the assessment of the systemic qualities of studies, one must be guided by plausible criteria that can be easily reproduced and applied by everyone. It is wrong to assume a binary classification, such that some studies do not meet the requirements of systemic analysis in any way, while others do so completely. Rather, we are dealing with a continuum of more or less strong systemic traits (Figure 3).

Decisive criteria for the inclusion of a systemic perspective in an evaluation design should be, in particular, the positive response to following questions:

- Is the realm of reality in which the intervention takes place characterised by properties that are typical of complex adaptive systems such as complexity, non-linearity, self-organization and emergence?
- Are expectations regarding desired outcomes of the intervention characterised by a high degree of uncertainty?
- Are serious results of the intervention reasonably to be expected only after long periods of time and dependent on many external and internal factors of the system, including its self-organising processes, which cannot be controlled by the policy-maker?
- Can consideration of the different perspectives of the programme makers and stakeholders involved contribute significantly to a better understanding of the programme?
- Does the delineation of the boundaries of the object of evaluation raise questions that are relevant to the assessment of the intervention?

Figure 3: Characteristics of systemic and non-systemic evaluations approaches
Source: Own depiction.



7 CAUSES OF LOW RECEPTIVENESS TO SYSTEMS THINKING

As we discussed above, elements of systems thinking can certainly be found in the practical evaluation of innovation policy measures, but there is hardly any question of a broad reception and general use of corresponding research instruments in projects predestined for this purpose. Barriers to a stronger echo of systems and complexity research in the evaluation of innovation policy measures are to be found in both epistemological and institutional fields. This is a topic that would deserve a substantially more elaborate discussion than is possible here. Some probably important aspects are elaborated here.

The human mind is primarily calibrated to the perception of simple, linear causal chains and is inclined to reproduce the mental model of the perception of linear causalities that proves itself anew every day. This serves as a ubiquitous model of knowledge even where it is not or only partially suitable for understanding a situation and often leads to erroneous attribution of blame for developments that have complex causes. Moreover, the evolution of the development of the natural sciences in western societies since the scientific revolution has fostered a type of technical rationality that has reinforced the predominance of linear causal thinking to the detriment of a willingness to adopt holistic perspectives (Meadows 2008). Of course, systems and complexity theory itself is a result of the development of technical rationality and arose in the mid-20th century in an effort to solve complex practical problems using sophisticated mathematical methods.

It should be noted, however, that systemic thinking, although by no means completely alien to people, tends to lose out in everyday life in comparison to linear thought patterns for epistemological and cultural reasons (Beasley 2012; Dörner 1997; Meadows 2008: 4), which is equally true for lay and professional people. The observation that the approaches of most innovation policy measures in the industrialised countries still follow the linear model of the innovation process today, despite its abandonment by innovation research decades ago, is probably largely due to this. Against this background, the spread of systems thinking in evaluators' community, but also in state bureaucracies, and above all its active use in practical work, does not happen automatically.

Another epistemological factor lies in the nature of the basic discipline of systems analysis and the particularities of its application. Systems and complexity research uses complicated mathematical models and is therefore easily suspected of being a playground for the mathematically gifted, who are likely to be found among practical programme evaluators only to a limited extent. There is a fundamental misunderstanding here. There is obviously no necessity to make system evaluations of government interventions dependent on the development of sophisticated mathematical models of the object of evaluation. In most cases, such an attempt would not be justifiable in any way in terms of financial and human resources. Rather, as Arnold (2004) shows, the use of system dimensions can be designed very differently depending on the object of study and the evaluation context. Incidentally, Bonini's paradox applies[3]: The more one tries to reflect complex systems in mathematical models by capturing as many relevant elements and relationships as possible, the less suitable this increasingly realistic model is for understanding reality.

Another factor that may be important in explaining the low reception of the systems concept are the worldviews sometimes conveyed together with it. As justified as the call for taking into account different perspectives on the object of evaluation and the emphasis on the role of boundary setting are (e.g., Williams 2015), they easily expose themselves to suspicion of ideology in practical contexts. The fact that the recent unprecedented rise of systems thinking is closely linked to the climate policy debate (e.g., Ison and Straw 2020; Meadows 2008) does not necessarily make things easier.

In addition to epistemological barriers, institutional barriers must also be taken into account. The mostly hard departmental demarcation between the ministries and the internal structuring of tasks in the ministries into clearly defined areas does not necessarily make it easier to incorporate concepts that imply cross-ministerial and cross-departmental cooperation. Proposals from outsiders who want to shake up the established boundaries of the areas of responsibility are not necessarily received with enthusiasm. In this respect, political decision-makers' declarations of intent to take a more systemic approach in the future should be treated with caution.

State bureaucracies have an apparent ability to take unwelcome ideas on the surface and distort them beyond recognition in the political process. An example of this is the 50-year "struggle" of the different German federal governments against bureaucracy, documented in a multitude of political documents, which continues to flourish and thrive despite all efforts in this regard. The same could happen with systemic thinking. This argument is not about finding culprits - this would be an expression of linear causal thinking -, but about elementary mechanisms of the development of bureaucracies.

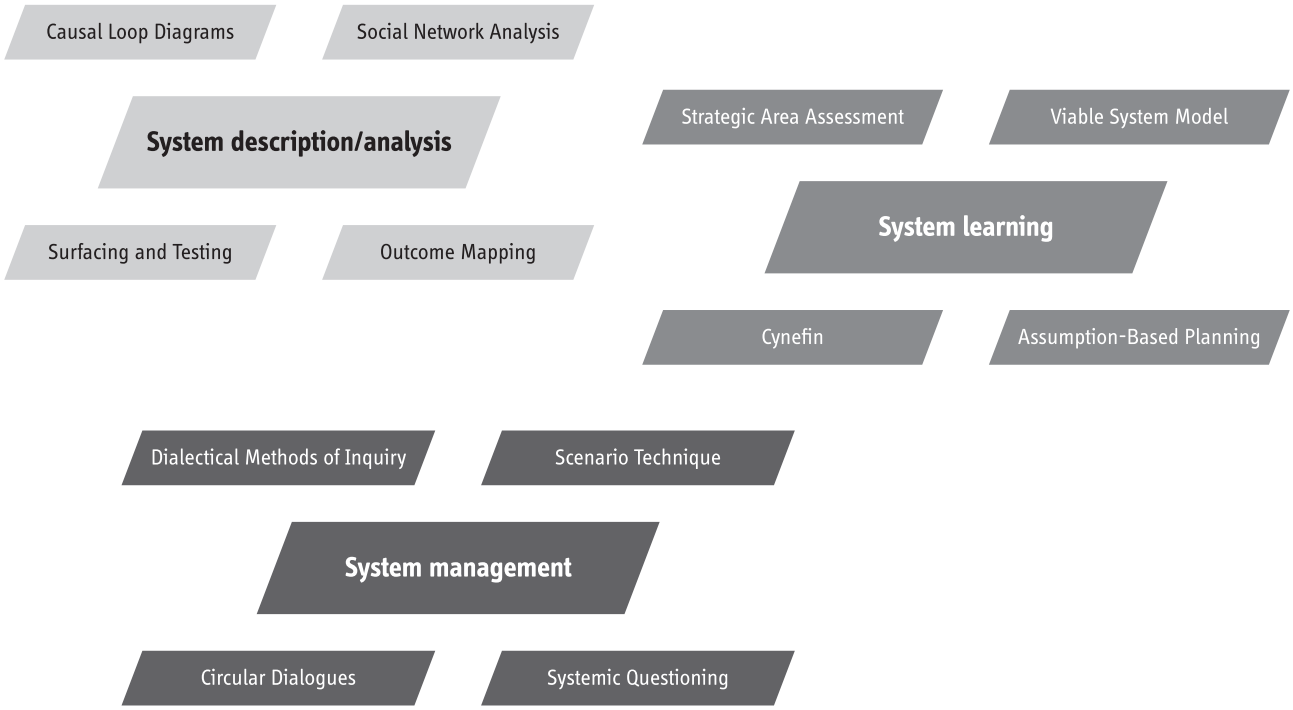
8 SUMMARY AND PROPOSALS FOR A PRAGMATIC USE OF SYSTEMS APPROACHES

Although the receptiveness for the inclusion of systemic perspectives in research has increased strongly in many scientific fields in recent decades, the potentials inherent in such approaches have so far only been used to a rudimentary extent in the practical evaluation of innovation policy programmes. This applies equally to assessments of innovation policy as a whole, the application of systemic perspectives to the simultaneous use of different, complementary policy instruments to achieve complex objectives, and to the evaluation of individual innovation policy programmes, such as the promotion of technology clusters, which target a complex object in order to pursue complex economic and technological

3 See wikipedia: „Bonini's paradox. Retrieved from https://en.wikipedia.org index.php?title=Bonini%s_paradox&oldid=1029627549, 21 June 2021.

objectives. Deficits in the first two mentioned areas have been repeatedly pointed out recently (e.g., Borrás and Laatsvit 2019; Edler and Fagerberg 2017). Our contribution aims to draw attention to the inclusion of systemic perspectives (a system-oriented framing) to the evaluation of individual programmes. This systemic framing can make useful contributions, and in many cases, it is only from this that a deep understanding of the state interventions in question can be generated.

Figure 4: The methodological arsenal of system-oriented evaluations



Cynefin is a knowledge management model developed by David Snowden and Cynthia Kurz for the analysis of complex adaptive systems. It is based on a typology of situations (simple, complicated, complex, chaotic), which takes a mediating position between complexity-reducing and complexity-emphasising procedures. The model takes into account the uncertainties inherent in complex adaptive systems that arise in the analysis and decision-making process (Williams and Hummelbrunner 2011: 163-183).

Source: Own depiction based on Williams and Hummelbrunner 2011.

Systemic evaluation amounts to a full consideration of the complexity properties of the object of study. It will always be part of a comprehensive methodological design that provides for the triangulation of the methodological tools contained in the mixed-method design. Therefore, under the umbrella of a systemic evaluation, rigorous quantitative methods such as RCTs will be found as well as qualitative components. What should count is solely the suitability of the chosen combination of methods for the best possible fulfilment of the research task at hand. Depending on the research task and the research context, the specific methods that have been treated in the evaluation literature as an expression of systemic research approaches should also be taken into account (Figure 4).

A systemic perspective is in general compatible with other prominent evaluation approaches, not least with realist evaluation and the diverse approaches of theory-based evaluation (Giel 2013, on the combination with complexity theory-based approaches Stame 2004). Both concepts have provided essential impulses for practical evaluation in multiple policy areas. Ray Pawson, pioneer of realist evaluation, seems to be sceptical of competing evaluation approaches that claim to do justice to the complexity of the evaluation object, including the systemic perspective (Pawson 2013: 53ff.). Other authors, however, point to the compatibility of systems/complexity approaches with realist evaluation (Westhorp 2012).

There are some practical steps that would lead practical evaluations to come closer to the idea of systemic evaluation:

i) Tailoring an evaluation programme that fits the characteristics of the object of evaluation in the best possible way;
ii) using instruments of complexity research in a complementary and supplementary function in normal routine evaluations of systemic evaluation objects;
iii) building experience driven models of the object of investigation without the use of overly complex-theoretical and mathematical constructions; this also means taking into account the complexity of interdependencies, uncertainties and emergent processes that lead to results of innovation funding;
iv) combining a system-oriented framing of an evaluation with all conventional (quantitative and qualitative) evaluation methods.

A systems approach can prove useful even in the case of rather simple innovation programmes (like "go-cluster").

Although for logical-systematic reasons there can be no evaluation that does not address systemic aspects in some way, the comprehensive consideration of dynamic systemic relationships in innovation policy evaluations has so far remained an exception. As we have tried to show, both epistemological and institutional factors are responsible for the unwillingness to adopt a systemic perspective in innovation policy evaluations.

It seems likely that evaluators of complex innovation programmes will claim that they were aware of the complexity of their object of study in the work process. They furthermore would make practical efforts to do justice to this complexity in the construction of their method design as well as in the practical evaluation work, at least insofar as the practical circumstances permit this. The lack of access to relevant data, the limited resources available and the time constraints of the evaluation alone would not have allowed this. An uneasy feeling might remain that essential things about the object of evaluation have not been revealed in the resulting research reports. Arnold (2004) has demonstrated that systemic perspectives can be introduced into such analyses even under conditions of numerous restrictions.

REFERENCES

Ackoff, R.L. (1971) Towards a System of Systems Concepts. Management Science 17(11): 661-671.

AEA (2015) Evaluating Outcomes of Publicly-Funded Research, Technology and Development Programs: Recommendations for Improving Current Practice. Version 1.0. Prepared by the Research, Technology and Development Evaluation Topical Interest Group of the American Evaluation Association (AEA), February 2015. Download under www.eval.org 21 January 2020.

Arnold, E. (2004) Evaluating research and innovation policy: a systems world needs systems evaluations. Research Evaluation (13)1: 3-17. Download from https://academic.oup.com/rev/article-abstract/13/1/3/1593255, on 24 February 2020.

Barbrook-Johnson, P., B. Castellani, D. Hills, A. Penn and N. Gilbert (2021) Policy evaluation for a complex world: Practical methods and reflections from the UK Centre for the Evaluation of Complexity across the Nexus. Evaluation 27(1): 4-17.

Beasley, R. (2012) 4.3.1 The Barriers to Systems Thinking. INCOSE International Symposium – July 2012. DOI: 10-1002/j.2334-5837.2012.tb01353.x.

Borrás, S. and J. Edler (2020), The roles of the state in the governance of socio-technical systems' transformation. Research Policy 49(5) 103971: 1-9.

Borrás, S. and C. Edquist (2019) Holistic Innovation Policy. Theoretical Foundations, Policy Problems, and Instrumental Choices. Oxford: Oxford University Press.

Borrás, S. and M. Laatsit (2019) Towards system oriented innovation policy evaluation? Evidence from EU28 member states. Research Policy 48(1): 312-321.

Bush, V. (1960) [1945]. Science – the endless frontier. A report to the President on a Program for Postwar Scientific Research. July 1945. Reprinted July 1960. Washington, D.C.: National Science Foundation.

Chaminade, C., B-A. Lundvall and S. Heneef (2018) Advanced Introduction to National Innovation Systems. Elgar Advanced Introductions. Cheltenham, UK, Northampton, USA: Edward Elgar.

Deutscher Bundestag, 19. Wahlperiode (2019) Bericht über die Programme zur Innovations- und Technologieförderung im Mittelstand in der laufenden Legislaturperiode, insbesondere über die Entwicklung des Zentralen Innovationsprogramms Mittelstand (ZIM). Fortschrittsbericht für das Jahr 2018. Unterrichtung durch die Bundesregierung. Drucksache 19/14480. Berlin: Deutscher Bundestag.

Dörner, D. (1997) [1089] Die Logik des Mißlingens. Strategisches Denken in komplexen Situationen. Rororo science. Reprint of paperback edition. Reinbek bei Hamburg: Rowohlt. Taschenbuch Verlag.

Easterly, W. (2007) [2006] The White Man's Burden. Why the west's efforts to aid the rest have done so much ill and so little good. Paperback edition. Oxford: University Press.

Edler, J., J. Fagerberg (2017) Innovation policy: what, why, and how. Oxford Review of Economic Policy. 33(1): 2-23. Download from: https://academic.oup.com/oxrep/article-abstract/33/1/2/2972712, on 19 November 2019.

Edquist, C. (2005) Systems of Innovation: Perspectives and Challenges. In J. Fagerberg, D.C. Mowey and R.R. Nelson (Eds.) The Oxford Handbook of Innovation. Oxford: Oxford University Press, pp. 181-208.

Eckert, S., V. Schüren and A. Bode (2016) Evaluation des Programms go-cluster des Bundesministeriums für Wirtschaft und Energie (BMWi). Abschlußbericht Berlin und Bickenbach: INTERVAL GmbH and CONABO UG (haftungsbeschränkt).

Forss, K., M. Marra and R. Schwartz (Eds.) (2017) [2011] Evaluating the Complex. Attribution, Contribution, and Beyond. Comparative Policy Evaluation 18. London and New York. Routledge.

Forss, K. and R. Schwartz (2017) [2011] Introduction. In Forss, K., M. Marra and R. Schwartz (Eds.) Evaluating the Complex. Attribution, Contribution, and Beyond. Comparative Policy Evaluation 18. London and New York. Routledge, pp. 1-32.

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

HM Treasury (2020) Magenta Book 2020. Supplementary Guide: Handling Complexity in Policy Evaluation. London: HM Treasury.

Holland. J.H. (2014) Complexity. A Very Short Introduction. Oxford: Oxford University Press.

Hummelbrunner, R. (2011) Systems thinking and evaluation. Evaluation 17(4): 395-403, DOI:10.1177/1356389011421935.

Imam, I., A. LaGoy and B. Williams (2006) Introduction. In Williams, B. and I. Imam (eds.) (2006) Systems Concepts in Evaluation: An Expert Anthology. American Evaluation Organisation, pp. 3-10.

Ison, R. and E. Straw (2020) The Hidden Power of System Thinking. Governance in a Climate Emergency. London and New York: Routledge.
Lundvall, B.A. (2010) National Systems of Innovation. Toward a Theory of Innovation and Interactive Learning. The Anthem Other Canon Series. London, New York, Delhi: Anthem Press.

Mainzer, K. (2008) Komplexität. UTB 3021. Paderborn: Wilhelm Fink.

Mainzer, K. (2014) Die Berechnung der Welt. Von der Weltformel zu Big Data. München: C.H.Beck.

Meadows, D.H: (2008) Thinking in Systems. A Primer. Edited by Diana Wright, Sustainability Institute. White River Junction, Vermont: Chelsea Green Publishing.

Midgeley; G. (2006) Systems Thinkling for Evaluation. In: B. Williams and I. Imam (Eds.) Systems Concepts in Evaluation: An Expert Anthology. American Evaluation Organisation

Mitchell, M. (2009) Complexity. A Guided Tour. Oxford, New York: Oxford University Press.

Mowles, C. (2014) Complex, but not quite complex enough: The turn to the complexity sciences in evaluation scholarship. Evaluation 20(2): 160-175.

Moyo, D. (2009) Dead Aid. Why aid is not working and how there is another way for Africa. London: Penguin.

Pawson, R (2013) The Science of Evaluation. A Realist Manifesto. Los Angeles et al.: Sage.

Rothgang, M., M. Peistrup and B. Lageman (2011) Industrial Collective Research Networks in Germany: Structure, Firm Involvement and use of Results. Industry and Innovation 18(4) 393-414.
RWI and WSF (2010) Erweiterte Erfolgskontrolle beim Programm zur Förderung der IGF im Zeitraum 2005 bis 2009. Endbericht. Forschungsvorhaben des Bundesministeriums für Wirtschaft und Technologie. Essen: RWI und WSF.

Schot, J. and Steinmueller, W.E. (2018). Three frames for innovation policy. R&D, systems of innovation and transformative change. Research Policy 47: pp. 1554-1567.Smits, R. and S. Kuhlmann (2004) The rise of systemic instruments in innovation policy. International Journal of Foresight and Innovation Policy. 1(1/2): 4-32.

Stame, N (2004) Theory-based Evaluation and Types of Complexity. Evaluation 10(1): 58-76.

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

Williams, B. (2015) Prosaic or Profound? The Adoption of Systemic Ideas by Impact Evaluation. IDS Bulletin 46(1): 7-16.

Williams, B. and R. Hummelbrunner (2011) Systems Concepts in Action: A Practitioner's Toolkit. Stanford, CA: Stanford University Press.

Williams, B. and I. Imam (eds.) (2006) Systems Concepts in Evaluation: An Expert Anthology. American Evaluation Organisation.

AUTHORS

MICHAEL ROTHGANG
RWI - Leibniz Institute for Economic Research
Hohenzollernstr. 1-3, 45128 Essen, Germany
E: michael.rothgang@rwi-essen.de
ORCID ID: 0000-0001-8328-0295

BERNHARD LAGEMAN
RWI - Leibniz Institute for Economic Research
Hohenzollernstr. 1-3, 45128 Essen, Germany
E: bernhard.lageman@t-online.de

KEYWORDS: Systems evaluation; innovation policy; systems analysis; complexity; innovation policy evaluation