THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

# Transformative innovation policy evaluation: characteristics, challenges, and lessons from practice

CAROLINA RESENDE HADDAD

Department of Technology Management and Economics

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2023

“You know my methods, Watson. There was not one of them which I did not apply to the inquiry. And it ended by my discovering traces, but very different ones from those which I had expected.”

- The Memoirs of Sherlock Holmes (1893)

Sherlock Holmes in “The Crooked Man” (Doubleday p. 416)

# Abstract

The emergence of the transformative innovation policy (TIP) framing, which builds upon previous science, technology, and innovation policy (STI) approaches to address transformative change or "grand challenges," implies new challenges for policy evaluation. While most frameworks for TIP evaluation are quite similar to those of previous STI policy evaluation approaches, it is here argued that issues such as capturing the effect of such policies, such as related to the direction of transformative change and its additionality, remains unaddressed in the literature. This PhD thesis focuses on three research questions: (RQ1) From a theoretical perspective, what could be the distinguishing characteristics of an "ideal type" TIP evaluation compared with previous innovation policy evaluation approaches? (RQ2) How can additionality and directionality be accounted for in TIP program evaluation? (RQ3) What are the challenges related to translating an "ideal type" TIP evaluation approach into practice?

RQ1 is mainly addressed by insights provided by Papers I and II. Specifically, Paper I develops a systematic literature review of TIP and identifies challenges for TIP evaluation, while Paper II develops an analytical framework specifying how key evaluation categories differ for an "ideal type" TIP evaluation compared with previous STI approaches. RQ2 is covered mostly by Papers III and IV. Paper III illustrates the development of the framework, which combines insights from theory-based evaluation and TIP to address directionality and additionality. Paper IV complements Paper III by specifying key transformative processes evaluators should look for when evaluating the effects of TIP interventions. RQ3 in turn is covered mainly by Papers II and V. Paper II uses the aforementioned analytical framework to make a cross-case analysis of three Swedish innovation programs funded by Vinnova, the Swedish Innovation Agency: Vinnväxt, Challenge-Driven Innovation (CDI), and the Strategic Innovation Programs (SIPs). Based on this analysis, it identifies three challenges for making evaluations more aligned with an "ideal type" TIP evaluation. Paper V critically analyzes the feasibility of the integrated framework from Paper III by performing an evaluation of the BioInnovation SIP to understand the challenges that emerge for evaluation practice. Finally, the thesis discusses the key characteristics of an "ideal type" TIP evaluation, reflects on the feasibility of the integrated framework for TIP program evaluation, highlights remaining issues for TIP evaluation, and concludes by presenting the key contributions of the thesis.



**Keywords**: Transformative innovation policy; Policy evaluation; Directionality; Additionality.

# List of appended papers

**Paper I**

Haddad, C.R., Nakić, V., Bergek, A., Hellsmark, H., 2022. Transformative innovation policy: A systematic review. *Environmental Innovation and Societal Transitions*, 43, 14-40.

**Paper II**

Haddad, C.R., Wise, E., Arnold, E., 2023. When theory meets practice in transformative innovation policy evaluation: Experiences from Sweden. Paper submitted for an academic journal.

**Paper III**

Haddad, C.R., Bergek, A., 2023. Towards an integrated framework for evaluating transformative innovation policy. *Research Policy*, 52(2), 10466.

**Paper IV**

Bergek, A., Haddad, C.R., 2022. Evaluating transformative innovation policy outcomes as unfolding processes of change in sociotechnical configurations, in: Vélez-Cuartas, G., Romero-Goyeneche, O.Y. (Eds.), *Transformative Metrics: Contributions to the Studies for Monitoring and Evaluating How Science, Technology, and Innovation Can Address Social and Environmental Challenges*. Universidad de Antioquia, Medellín.

**Paper V**

Haddad, C.R., 2023. Identifying challenges for the evaluation of transformative innovation policy programmes: insights from the BioInnovation SIP in Sweden. Paper submitted for an academic journal.

## Other related work

Haddad, C.R. (2019). A functions approach to analyze CE systemic performance in EU28 countries. International Sustainability Transitions Conference 2019, June 23-26, Ottawa, Canada.

Haddad, C. R., Nakic, V., Bergek, A., Hellsmark, H. (2019). The policymaking process of transformative innovation policy: a systematic review. 4th International Conference on Public Policy, June 26-28, Montréal, Canada.

Haddad, C.R., Bergek, A. (2020). A functions approach for evaluating transformative innovation policy. International Sustainability Transitions Conference 2020, August 18-21, online.

Haddad, C.R., Volan, T., Uriona-Maldonado, M., Vaz, C. (2020). The Electric Vehicle diffusion in Sweden: will the country achieve its zero-emission fleet by 2030? System Dynamics Conference 2020, July 19-24, online.

Bergek, A., Haddad, C.R. (2020). A functions approach for evaluating transformative innovation policy: indicators and metrics. Transformative Metrics Workshop, October 5-6, online.

Haddad, C.R. (2021). Evaluating transformative innovation policy: a systems analysis of the BioInnovation Strategic Innovation Program in Sweden. International Sustainability Transitions Conference 2021, October 5-8, online.

Keenan, M., Haddad, C.R. (2022). S&T Policy 2025: Rethinking STI policy for transitions. EU-SPRI Conference 2022, June 1-3, Utrecht, The Netherlands.

Haddad, C.R., Bergek, A. (2022). Rethinking additionality for transformative innovation policy. EU-SPRI Conference 2022, June 1-3, Utrecht, The Netherlands.

Haddad, C.R., Wise, E., Arnold, E. (2023). When theory meets practice in transformative innovation policy evaluation: experiences from Sweden. EU-SPRI Conference 2023, June 14-16, Brighton, England.

# Acknowledgements

First and foremost, I would like to express my gratitude to Anna Bergek for being an incredible supervisor throughout the completion of my PhD studies. I really appreciate all your support, guidance, kindness, and patience. Thank you for allowing me to develop my own ideas and for consistently challenging them. I have learned so much from you in these past years and it was an honor to be your PhD student.

I also want to thank Hans Hellsmark for being part of my research team as a co-supervisor. Thank you for questioning my ideas, for challenging me to improve my research, and for making our meetings enjoyable.

A big thank you to all my colleagues from the Swedish Strategic Innovation Program Platform (STIPP) and Vinnova for allowing the development of such a challenging and inspiring research project.

To my colleagues at the Environmental Systems Analysis (ESA) division, thank you for always being supportive and for encouraging me to do my best, both as a researcher and as a person. A special thanks to the ESA collective gang, some of whom have also become friends for life (you know who you are). I really appreciate all our laughs, lunches, dinners, etc. You really made my PhD life easier and more enjoyable.

I would also want to thank my colleagues at the Governing Resilience and Transformation (GReaTr) project. A special thanks to Emily Wise and Erik Arnold. It was a pleasure to co-author a paper with you. I learned a lot from you and really enjoyed our “nerd” talks throughout the process.

A special thanks to Michael Keenan for guiding and supporting me during my internship at the OECD and to my other colleagues who were very supportive during my time in Paris. I really enjoyed working with you.

A big thanks to Diego Hernando Florez Ayala, my transitions colleague and friend of many years. Thank you for all transitions talks and for always being there for me when I needed during my studies. Thank you also for putting a lot of effort into creating my thesis cover.

I am also really grateful to have such a wonderful family: My dear parents and stepfather, who have always believed in and supported me. My dear grandparents, who have always been there for me and are my great role models in life. I lost two of you during my PhD studies, but I cannot express in words how grateful I am to have had you in my life. My brother, who has always been on my side, believing, supporting, and encouraging me in every step. My sister, who is always teaching me something new on social media. I love you all and miss you every day.

A special thank you to my husband, Gustavo, for always being there for me. Thank you for supporting my decisions, for questioning me and for enduring me during this process. Thank you for being my PhD guide, my friend, and my partner. I love you.

# Table of Contents

# 1 Introduction

This PhD thesis focuses on transformative innovation policy evaluation. Transformative innovation policy (TIP) refers to a new framing of innovation policy that studies how policies can drive innovation toward socio-technical system transition. In this introductory section, I discuss TIP as an emerging framing of innovation policy in comparison with previous science, technology, and innovation (STI) policy framings. I also give a brief overview of how STI policy evaluation has evolved since the 1960s until new approaches emerged connected to TIP evaluation. These aspects form the background of this PhD thesis and set the context of my research problems and questions.

Herein, the word "policy" is used as an umbrella term that covers policy goals as well as policy instruments at various levels of aggregation (projects, programs, larger initiatives or instrument mixes). The word "evaluation" refers to the process of examining the intended and actual effects of public policies to generate insights regarding policy outcomes and impacts. Both these concepts will be explored in more detail in Section 2.

## 1.1 Background

### 1.1.1 Transformative innovation policy as a new framing[1]

In recent years, transformative innovation policy has emerged as a third innovation policy framing (Schot and Steinmueller, 2018, p. 597) or paradigm[2] (Diercks et al., 2019), which builds upon two established framings of innovation policy.

*Framing 1*, also called the science and technology policy paradigm, has been influential since World War II (Diercks et al., 2019). It views innovation as oriented toward economic growth and competitiveness. This framing was influenced mainly by neoclassical economics and later by new growth theory (Chaminade and Edquist, 2008). For neoclassical theorists, innovation is seen as a linear sequence of phases (Chaminade and Edquist, 2008), in which "one does research, research then leads to development, development to production, and production to market" (Kline and

[1] Adapted from Haddad (2021b, pp. 1-7).
[2] A policy paradigm is defined by Hall (1993, p. 279) as "a framework of ideas and standards that specifies not only the goals of policy and the kind of instruments that can be used to attain them, but also the very nature of the problems they are meant to be addressing."

Rosenberg, 1986, p. 285). The standard rationale for policy intervention is market failures: because scientific knowledge is uncertain, cannot be appropriated, and depends on previous investments (Arrow, 1962; Nelson, 1959), private actors fail to efficiently allocate resources to the extent that would be socially and economically desirable, and hence policymakers need to intervene (Chaminade and Edquist, 2008). Governments should intervene to fix markets by investing in areas portrayed by positive or negative externalities, information asymmetries, and capital market failures (Jacobsson et al., 2017; Mazzucato, 2016).

The early 1980s were characterized by the emergence of evolutionary economics and the notion of interactive innovation processes (Martin, 2012). This was driven, among other things, by the perceived limitations of neoclassical economics, and resulted in a broader view of innovation (Isaksen, 1999) as well as an increased focus on "opportunity enhancing" innovation policies (Georghiou, 1998). While this did not directly influence the rationales of innovation policy, it did increase attention to policies directed at increasing collaboration and technology transfer between firms and other actors. In the late 1980s, the evolutionary and interactive perspective on innovation converged with a revived discussion of the merits of industrial policy, which resulted in the development of various "innovation systems" frameworks and an increased focus on "systemic" innovation policies (Smits and Kuhlmann, 2004). This is referred to as *framing 2* of innovation policy (Schot and Steinmueller, 2018). In general, an innovation system involves the creation, diffusion, and use of knowledge and is formed by its components, as well as their relationships, characteristics, and attributes (Carlsson et al., 2002). Many approaches emerged aiming at understanding the innovation processes at the national (Freeman, 1987; Lundvall, 1992; Nelson, 1993), technological (Carlsson and Stankiewicz, 1991), sectoral (Malerba, 2002), and regional levels (Cooke et al., 1997).

These "systemic" framings of innovation policy brought a shift in rationales for government intervention from market failures to various system failures or weaknesses that may characterize spatially or cognitively delineated innovation systems (Chaminade and Edquist, 2010).[3] On the one hand, these include structural deficiencies related to capabilities, networks, infrastructure, and

[3] Spatially delineated systems include regional and national innovation systems. Cognitively delineated systems include technological and sectoral innovation systems. For a discussion of spatial and cognitive proximity, see Boschma (2005).

institutions (cf., e.g., Carlsson and Jacobsson, 1997; Jacobsson and Johnson, 2000; Tödtling and Trippl, 2005; Woolthuis et al., 2005). On the other hand, they include problems related to key innovation processes (or "functions") (Bergek et al., 2010; Wieczorek and Hekkert, 2012), which directly influence the development, diffusion, and use of new technologies such as market formation, resource mobilization, and entrepreneurial experimentation (Bergek, 2019; Bergek et al., 2008).

Building upon these previous innovation policy framings, *framing 3* has been labeled transformative innovation policy and "can be seen as layered upon, but not fully replacing" earlier framings of innovation policy (Diercks et al., 2019, p. 881). According to Diercks et al. (2019), TIP brings changes in both the policy agenda and the understanding of the innovation process (i.e., the activities contributing to the generation of innovations and the actors involved in the innovation process). Regarding *policy agenda*, transformative innovation policy goes beyond the idea of economic growth and conventional innovation systems and instead calls for directionality in innovation studies and the need to address "Grand Challenges." Additionally, it targets policy domains beyond traditional economic and industrial policies and pushes toward acknowledging the pro-innovation bias logic *vis-à-vis* the negative implications of new innovations.

In relation to the *innovation process*, TIP is often conceptualized in terms of socio-technical system change, as proposed by the broader sustainability transition community, which emphasizes that transformative change entails not only technological change, as fostered by established innovation policy, but also changes in user practices, markets, and institutions, among others (Diercks et al., 2019; Schot and Steinmueller, 2018).[4] TIP also involves multiple actors beyond the "triple helix" of university–industry–government relations (cf. Etzkowitz and Leydesdorff, 2000) to include societal actors, such as citizens. Moreover, TIP encompasses the promotion of both supply- and demand-side policies, as in the innovation system approach, and highlights the need to include policies that address the destabilization of incumbent regimes (Kivimaa and Kern, 2016).

[4] Hölscher et al. (2018) discuss the difference between transition and transformation, stating that transitions emphasize changes in societal subsystems, focusing on interactions between society, technology, and institutions, whereas transformation regards chances in whole societies (local, national, global), including the human and biophysical components of a system. The concepts are not mutually exclusive, but offer nuanced perspectives on how to explain non-linear societal change.

Orienting innovation policies toward addressing societal challenges also implies additional rationales for policy intervention. According to Weber and Rohracher (2012), policies for transformative change should draw not only on the market and system failures perspective, but also on transformational system failures. The authors thus propose a third category of failures centered around four rationales for transformative change, which integrates insights from innovation systems and transition approaches. These include directionality, demand articulation, policy coordination, and reflexivity failures. In particular, directionality refers to the need to build innovations that are as efficient and effective as possible, and also contribute to a direction of change. Given the emphasis on directionality in TIP, this becomes a key issue for evaluation and hence will be considered further in this thesis.

### 1.1.2 Evolution of STI policy and TIP evaluation[5]

The evolution of different innovation policy framings has also been reflected in the evaluation of science, technology, and innovation (STI) policy (Molas-Gallart and Davies, 2006). Although STI policy evolved later than other policy fields, such as education and psychology, whose evaluation practices had already been established by the end of the 1960s (Gök, 2010; Molas-Gallart and Davies, 2006), early STI policy evaluation studies can be traced back to the 1960s in the United States and the 1970s in Europe (cf. Luukkonen, 2002; Roessner, 2002). Internationally, the Organisation for Economic Co-operation and Development (OECD) took the lead in developing guidelines for R&D statistics in the form of the Frascati and Oslo manuals.

In 1963, the OECD released the first version of the Frascati Manual, which contained data collection guidelines for using R&D statistics (OECD, 2015). The manual was the first to standardize the practice of surveys of research and development (OECD, 2015), and served as a landmark for the collection of standardized statistics in different countries. From the 1980s until the mid-1990s, the OECD organized a series of conferences and workshops to discuss innovation measurement and indicators, which would lead to the development of the first version of the Oslo Manual (OECD, 1992). The Oslo Manual is now in its fourth edition and has constituted an

[5] Adapted from Haddad (2021b, pp. 7-13).

important step toward gathering information to investigate the innovative performance of countries using a more systemic view of innovation policy (Lundvall and Borrás, 2005).

In parallel with the activities of the OECD, the concept of *additionality* started to be developed and used in the UK in the early 1980s as a way to measure "the difference which government-sponsored programmes have made to the recipients, particularly companies, in terms of R&D activities" (Luukkonen, 2000, p. 711). Furthermore, it has been the hallmark of technology and innovation policy impact evaluation (Gök, 2010). While the concept of additionality seems to have evolved largely independently of developments of different innovation policy framings, its origins can be traced to the neoclassical market failure rationale (Luukkonen, 2000). Thus, changing policy objectives have also influenced the type of additionality policymakers were aiming at. For evaluation purposes, Georghiou (1994) differentiates between three types of additionality: input, output, and behavioral additionality.

*Input additionality* relates to the neoclassical perspective and refers to the extent to which public funding increases total R&D investments instead of replacing or crowding out firm's investments (Georghiou and Clarysse, 2006). *Output additionality* emerged in response to criticisms of using input additionality, which assumes that there is a direct linear link between R&D input and innovation output (Clarysse et al., 2009). It measures whether the same outputs, such as patents, publications, and sales, would have been achieved without public support (Bach and Matt, 2005; Clarysse et al., 2009; Georghiou and Clarysse, 2006). While these two types of additionality are useful measures of the performance of firms that received public support, they ignore changes in the behavior of firms and treat the firm as a black box (Georghiou and Clarysse, 2006). Therefore, *behavioral additionality* (BA), influenced by the evolutionary/structuralist perspective on innovation policy, emerged as a the measure of influence of a policy intervention in the actors of firms and other actors (Gök and Edler, 2012). Over time, the BA concept was expanded in an attempt to capture the indirect effects of policy, in particular in terms of the learning that takes place in firms because of public support (Clarysse et al., 2009).[6] A more detailed discussion of

[6] Autio et al. (2008) label behavioral additionality as "second-order additionalities" to refer to the broader range of learning effects emerging from the implementation of R&D subsidies, while input and output additionality are termed "first-order additionalities." According to Clarysse et al. (2009), different types of organizational learning include: experiential learning/learning-by-doing, congenital learning/absorptive capacity, and interorganizational learning/knowledge sharing between firms.

these three types of additionality can be found elsewhere (Haddad, 2021b). Given the importance of additionality in the impact evaluation of technology and innovation policy, I elaborate further on the concept later in this thesis.

With the uptake of the innovation system framing among scholars and policymakers has come an understanding that the effects of innovation policies should be evaluated at the system level rather than at the level of projects or firms, and some systemic approaches to this end were advanced (Arnold, 2004; Bellandi and Caloffi, 2010; Gök, 2010; Rametsteiner and Weiss, 2006; Russo and Rossi, 2009). For example, some authors focused on meta-evaluation or secondary analysis, combining evaluations at different levels beyond individual policy interventions, to capture the systemic effects of policy (Arnold, 2004; Edler et al., 2008). Others suggested the study of how innovation network composition and interactions change as a result of policy intervention, for example in terms of network composition, organization, and cohesion or the relative centrality of different actors (Bellandi and Caloffi, 2010; Rametsteiner and Weiss, 2006; Russo and Rossi, 2009). Yet others focused on analyzing structural system failures and assessing to what extent a specific policy instrument addresses these failures (Woolthuis et al., 2005), or investigating how policies influence key processes (or "functions") in technological innovation systems (TISs) (Bergek, 2004; Jacobsson and Perez Vico, 2010; Perez Vico and Jacobsson, 2012). Additionally, some authors highlighted the need to evaluate policy mixes rather than programs or meta-evaluations (Flanagan et al., 2011; Magro and Wilson, 2013).

To reflect the emergence of the TIP framing, several authors highlighted the need to develop new frameworks that capture more of the complex, system-level transition dynamic to allow evaluators to understand what is happening in the focal socio-technical system and how policy influences the conditions for realizing transformative change (Amanatidou et al., 2014; Hoppmann et al., 2014; Janssen, 2019; Kern, 2012; Mazzucato, 2016; Mazzucato, 2018). To reflect that, some new approaches for policy analysis and evaluation have been proposed, which combine insights from innovation and transition studies and draw on frameworks introduced for other purposes.[7] Some authors focus on policy failures, drawing either on Klein Woolthuis et al.'s (2005) systemic

[7] Other frameworks and approaches to evaluation have been developed for, among others, sustainable innovation policies, sustainable transition policies, and sustainable transformations (Boni et al., 2019). Here, I highlight those that target TIP specifically, as delineated in Section 1.1.1, as this is the focus of this PhD thesis.

failures framework van Mierlo et al. (2010) or on Weber and Rohracher's (2012) transformational failures framework (Bugge et al., 2017b; Bugge et al., 2018; Grillitsch et al., 2019). Another group of authors departs from the TIS functions to assess the design and impact of transformative policy (Janssen, 2019) or to analyze the influence of policy mixes toward low-carbon systems (Kivimaa and Virkamäki, 2014). Kivimaa and Kern (2016) combine the TIS functions with elements of "regime destabilization," which they argue are needed to better capture the destruction of incumbent regimes. Their framework has later been used to assess the impact of different types of policy mixes (Kivimaa et al., 2017; Scordato et al., 2018).

Other works are based more directly on transition-oriented frameworks. For example, Kern (2012) uses the MLP to make an *ex-ante* analysis of the impact of policies on processes at the niche and regime levels. Meelen and Farla (2013) combine insights from technology management (TM), SNM, and the TIS functions to propose an integrated framework for analyzing sustainable innovation policy interventions. Moreover, Ghosh et al. (2021) outline a set of 12 transformative outcomes (TO) related to key transition processes that can be used to guide the evaluation and reformulation of innovation policies targeting transformative change.

In this section, I briefly introduced the evolution of STI policy evaluation, from measurements of R&D statistics and country innovative performance to the emergence of different types of additionalities, and the application of failures and innovation system approaches for analyzing policy. Additionally, I have summarized some emerging attempts to evaluate TIP, which mostly adapt innovation and transition frameworks for analyzing transformative change. Despite these attempts, TIP evaluation remains an emerging topic and some challenges remain, as will be discussed in the next section.

## 1.2 Research problems and research questions[8]

As discussed in Section 1.1.2, the uptake of the innovation system approach led to the development of new frameworks for the analysis and evaluation of innovation policies to capture their complex and systemic characteristics. However, despite these advances, some authors have pointed out that innovation policy practice still lagged behind advances in innovation theory (Molas-Gallart and

[8] Partially adapted from Haddad (2021b, pp. 16-17).

Davies, 2006) and that only a few countries had been able to develop system-oriented innovation policy evaluation (Borrás and Laatsit, 2019). The emergence of TIP seems to be following a similar path. As described in Section 1.1.2, the rise of TIP was also followed by new attempts to analyze and evaluate policies targeting transformative change by adapting the transformational failures or current sustainability transitions frameworks to capture transformative change. While these frameworks discuss many ideas and point out some aspects that TIP evaluation should cover, to the best of my knowledge, to date there is no existing TIP-oriented evaluation practice. This might not only be because different agencies are still experimenting with TIP in different formats (Borrás and Schwaag Serger, 2022; Larrue, 2021), but also because it is unclear what an "ideal type"[9] TIP evaluation would encompass.

While I will return to the practical issues regarding TIP evaluation later in this section, it is first important to understand what characterizes an "ideal type" TIP evaluation, as it remains unclear what, in theory, distinguishes TIP evaluation from the evaluation approaches developed within previous innovation policy framings. Understanding this in greater detail is an important first step in capturing the main shortcomings of existing evaluation frameworks in relation to TIP evaluation and identifying issues that still need to be addressed. Accordingly, the first research question addressed by this PhD thesis is as follows:

**Research question 1:** From a theoretical perspective, what are the distinguishing characteristics of an "ideal type" TIP evaluation compared with previous innovation policy evaluation approaches?

What distinguishes TIP evaluation from previous innovation policy framings remains unclear and given my interest in evaluating the effects of policies targeting transformative change, two key challenges have been highlighted in the literature on TIP evaluation: the issues around directionality and additionality. Regarding the former, TIP evaluation involves challenges related to the increased focus on *directionality*, as TIP implies that there are new ends and goals against which policy interventions need to be evaluated. Instead of treating all innovation outcomes as positive, as under previous paradigms, there is a need to consider how policy interventions

[9] An "ideal type" refers to "a notion that defines the general traits of the expected phenomena, and which is used for analytical purposes" (Borrás and Laatsit, 2019, p. 314).

contribute to the ultimate goal of a policy in terms of the desired transition pathway(s) for a particular sector (Magro and Wilson, 2019; Schlaile et al., 2017), for example in terms of which actors are involved and which types of innovations are developed and diffused in the system (Geels et al., 2016). This implies asking neglected questions such as "which way?"; "who says?"; and "why," and not only "yes or no?"; "how much?"; "how fast?"; and "where?" (Andersson et al., 2021; Stirling, 2009).

Existing TIP frameworks, such as those discussed in Section 1.1.2, make some attempts to address directionality for policy analysis. For example, some researchers study directionality as the capacity to build a shared vision (Bugge et al., 2017a; Bugge et al., 2018; Scordato et al., 2018), others investigate the directionality challenges emerging from actors' interests and capabilities, networks, and institutions (Grillitsch et al., 2019), and others still see directionality as a result of the alignment of actors and the adaptations in policy interventions in relation to the long-term target of TIP (Janssen, 2019). However, most scholars do not discuss explicitly how to capture and assess directionality in TIP evaluation.

With regard to *additionality*, criticisms related to input and output additionality led to the development of the concept of behavioral additionality (BA) to capture changes in the behavior of firms due to policy. However, BA has also been the target of criticism. Specifically, Gök (2010) argues that current BA analysis still focuses on the behavior of individual firms without analyzing the implications of changes in the behavior of a population of firms and the long-term institutionalized routines within the economy. Additionally, despite being a popular concept in innovation policy evaluation, the application of BA still faces operational and methodological problems, such as the lack of a consistent unit of analysis and the use of comparative statistics as the exclusive method for analyzing BA (Amanatidou et al., 2014; Gök and Edler, 2012; Kubera, 2021).

If there were already methodological difficulties in establishing causal relationships and assessing the outcomes of policy over a long time in systemic evaluations (Purkus and Lüdtke, 2020), the emergence of TIP makes things even more complex. In TIP evaluation, first, a broader set of impacts on targeted systems than in previous evaluation approaches (Amanatidou et al., 2014) should be considered. Additionality assessments therefore need to go beyond evaluating the traditional economic, technological, and scientific spheres usually addressed by input and output

additionality (Amanatidou et al., 2014) to further address the societal impact of the policy measure and its long-term effects on the behavior of actors, as well as learning at the system level. Second, TIP involves even more complex feedback loops between policy outputs, outcomes, and impacts that should be accounted for (Arnold et al., 2018; Kern and Rogge, 2018). Therefore, current successionist approaches used for assessing BA are unsuitable as they do not fully capture contextual features and what happens between the cause and effect (Kubera, 2021).

Because of these challenges, additionality is rarely discussed in the frameworks targeting TIP evaluation, perhaps because of the focus on evaluation as a tool for policy learning rather than accountability (Amanatidou et al., 2014; Magro and Wilson, 2019). One exception is Janssen (2019, p. 83), whose impact framework to some extent allows for the analysis of additionality (in addition to output) by "disentangling the contribution of policy efforts to the strengthening of technological innovation systems." Thus, the question of how to operationalize BA in a TIP context has been inadequately addressed.

The need to address additionality and directionality in TIP policy thus leads to the second research question addressed in this thesis:

**Research question 2:** How can additionality and directionality be accounted for in TIP program evaluation?

While the aforementioned gaps are more conceptual and methodological, TIP evaluation also faces gaps related to the practice of evaluation. Previous studies of innovation policy evaluation practices identified mismatches between theoretical understandings of innovation policy evaluation and evaluation practices. For instance, Edler et al. (2012) analyzed the practice of innovation policy evaluation across Europe considering key evaluation dimensions, such as timing, the purpose of evaluation, and methods of data collection and analysis. The authors concluded that although these evaluations are both summative and formative, it is mainly the former that informs decision-making and adjustments of policies. Additionally, these evaluations only look at a limited range of impacts, primarily technological and economic ones, and primarily consider supply-side instruments, indicating that they are not fully adapted to capture systemic effects. Another example is from Gök and Edler (2012), who studied how behavioural additionality has been operationalized in evaluation practice in 25 European countries. The authors found that the use of BA in

evaluations is still limited, as evaluations do not consider the multiple dimensions of behavioral change in innovation performance and mainly look at the firm level. This indicates that if BA was to reflect the evolutionary perspective on innovation policy, it would not fulfil its function in evaluation practice.

More recently, Borrás and Laatsit (2019) looked for evidence in EU28 countries (pre-Brexit) regarding the evaluation practices toward an "ideal type" system-oriented innovation policies, considering dimensions such as: the coverage of the evaluation (e.g., in relation to assessing the entire policy mix), the inclusion of a systemic perspective (in terms its performance), the frequency of the evaluations (i.e., temporality), and diversity of expertise used in the evaluation. This latter work concluded that only a few countries seem to follow a systems approach to evaluation, even if many years have passed since the emergence of framing 2.

As highlighted by some authors, TIP theory seems to be somewhat "lost in translation" regarding how overall policy agendas are translated into policy practices (Bergek et al., 2023; Diercks, 2018; Ulmanen et al., 2022). As discussed above, this is also true for evaluation. Even though many years have passed since the emergence of framing 2, systemic evaluation has not entirely been embraced in practice, and the emergence of TIP brings additional challenges. However, research on the practice of TIP evaluation remains limited. One exception is Rohracher et al. (2022), who analyze the shift in discourse and practice of innovation policy toward transformation in Sweden and how this has affected program monitoring and evaluation. While this study provides important insights regarding the practice of program implementation and evaluation in the country, little is known about the extent to which policy practice has incorporated TIP thinking, operationalized in terms of the defining characteristics identified in RQ1, and the challenges that evaluators might face when trying to apply an "ideal type" TIP evaluation approach. This brings us to the third research question addressed in this thesis:

**Research question 3:** What are the challenges related to translating an "ideal type" TIP evaluation approach to practice?

## 1.3 Structure of this PhD thesis

In Section 1.1, I first outlined the background around the emergence of TIP and how STI policy evaluation evolved until new frameworks were proposed for evaluating TIP. This provided the

basis for identifying the research problems and research questions, as discussed in Section 1.2. In Sections 2 and 1, I outline the different approaches put forward in policy evaluation and discuss transformative innovation policy and the key theoretical frameworks that underlie this thesis. Section 4 then outlines my research methodology by contextualizing my research in relation to my research project and providing an overview of my research process, research design, and methodological limitations. In Section 5, I introduce the papers that are appended to the thesis and serve as the basis for addressing the research questions outlined in Section 1. Section 6 outlines my findings in relation to these research questions, while Section 7 discusses these findings. Finally, Section 8 concludes the main findings of the thesis, highlights my key contributions, and summarizes the main limitations and recommendations for future research.

# 2 Policy evaluation

In this section I explore the concepts of "policy" and "evaluation" in details. I thus start by defining these concepts, followed by a discussion of different perspectives on evaluation. As an inherently practical area, the evaluation literature is mainly composed of guiding frameworks that discuss how evaluations should be developed (Alkin, 2013; Knill and Tosun, 2012). I introduce different perspectives that have developed based on different epistemological discussions over time. This discussion helps me identify the most appropriate methodological starting point for evaluating TIP. This exercise is needed given that evaluation is a contested field, immersed in a compelling debate about the most appropriate epistemologies and political and methodological starting points (Rolfe, 2016). Therefore, this section will form the basis of my methodological choices, which will be covered in Section 4.

## 2.1 Defining public policy and evaluation

In Section 1, I stated that in this PhD thesis, I use "policy" as an umbrella term that covers policy goals and instruments at various levels of aggregation, that is, projects, programs, and larger initiatives or instrument mixes. In the public policy literature, there is no common definition of what public policy means, but rather authors adopt different definitions highlighting multiple aspects that arise within the term (Cairney, 2012). For example, public policies are "deliberate decisions by governments or equivalent authorities to take action or nonaction toward specific objectives" (Weible and Carter, 2017, p. 24). Taking action or nonaction means that policymaking revolves around a fundamental decision by the government to do or not do something regarding a problem. According to Howlett and Cashore (2020, p. 13), public policies are composed of various parts, that is, they are formed by goals and means that are combined and implemented by different policy actors who operate in a wider environment formed by other actors and institutions.

One way of conceptualizing public policies is through the policy mix concept. The term "policy mix" is used in different scientific fields, such as environmental economics (Lehmann, 2012), innovation policy (Flanagan et al., 2011) and policy analysis (Howlett and Rayner, 2007). The concept was picked up at the beginning of the 2000s by the analysists and scholars to refer to the "interactions and interdependencies between different policies as they affect the extent to which intended policy outcomes are achieved" (Flanagan et al., 2011, p. 702).

Later, Rogge and Reichardt (2016) extended the concept and conceptualized policy mixes as being composed of three main building blocks, as illustrated in Figure 1. According to the authors, the first building block, *elements*, refers to the content of the policy mix and consists of the policy strategy and instrument mix. The second block, *policy process*, refers to the policymaking process, which can include all the stages of the policy cycle. In general, a policy cycle is composed of different stages, such as agenda-setting, policy formulation, legitimation, implementation, monitoring and evaluation, and policy learning (Cairney, 2012; Howlett and Giest, 2013). In this cycle, evaluation is one of the later stages of the policymaking process. The policy *characteristics* in turn include the consistency of elements, coherence of processes, credibility, and comprehensiveness. These building blocks can also be specified along different dimensions, such as policy field, governance level, geography, and time, and they also relate to each other to promote technological change (Rogge and Reichardt, 2016).

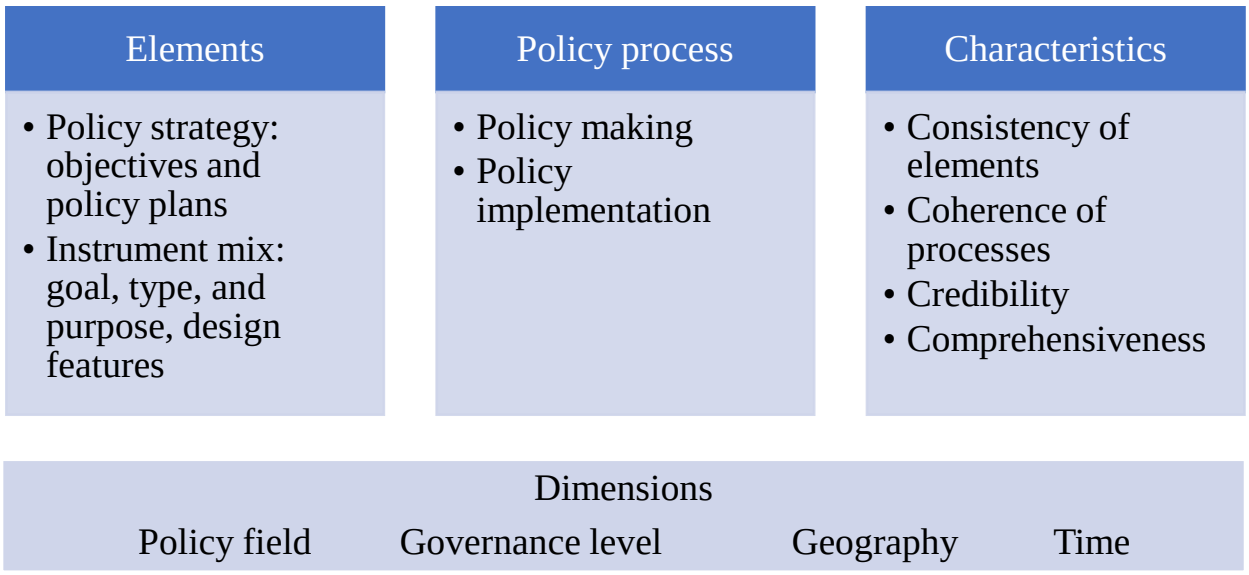


Figure 1. Policy mix building blocks. Source: adapted from Rogge and Reichardt (2016).

In this conceptualization, two building blocks are mostly relevant in this PhD thesis: the policy process and the elements of the policy mix. Although the characteristics of the policy mix are also important, I do not consider them directly in this thesis. Rather, I focus on the elements of the policy mix, mostly looking at the level of programs and, in particular, on the evaluation stage of the policy process. This is because, empirically, this thesis will consider current innovation policy programs that have been re-oriented over time to address societal challenges.

In Section 1, I stated that evaluation refers to the process of examining the intended and actual effects of public policies to generate insights regarding policy outcomes and impacts (Knill and Tosun, 2012). Similar definitions are provided by Weiss (1993), who states that evaluation looks at the effect of policies on the target population in relation to the goals they are set to achieve. She also emphasizes that evaluation is a political activity, wherein evaluation findings affect and are affected by the political context. As a political activity, evaluations are "socially constructed and politically articulated," as it is influenced by those being evaluated and are also used as a tool to provide meaning within social and political relations (Taylor and Balloch, 2005, p. 1). Lastly, evaluation can also be seen as a governance and learning tool used by policymakers to generate lessons for further policy practices (Amanatidou et al., 2014; Flanagan et al., 2011).

Over the years, different evaluation types and approaches have emerged and many attempts to group and categorize the various perspectives have been made (Alkin, 2013; Edelenbos and Van Buuren, 2005; Stufflebeam, 2001). Policy evaluation can be performed in several ways and for different reasons. Evaluations can be *formative*, for improving a policy measure by continuously supporting learning and decision-making processes, or *summative*, aiming at assessing policy impact by establishing cause-effect relationships (usually at the end of policy implementation) (Knill and Tosun, 2012). Patton (2010, p. 1) proposes a third approach, namely developmental evaluation, which "supports *development* to guide adaptation to emergent and dynamic realities in complex environments." According to the author, it offers an alternative to formative evaluation, which focuses on program improvement, to focus on program development, where the evaluator is often part of the development team, engaged in the design and continuous development of the intervention. Evaluations can also be conducted *ex-ante*, that is, before an intervention, or *ex-post*, that is, after the program was implemented (Khandker et al., 2010).

The literature also distinguishes between complementary approaches to evaluation, including monitoring, process evaluation, and impact evaluation (Gertler et al., 2011), which can support "evidence-based" policy (Pawson, 2006).[10] Monitoring is a continuous process of tracking inputs, activities, outputs, and, occasionally, outcomes to inform program implementation and daily

[10] While the term "evidence-based" policy can have a rather rational connotation, I agree with Pawson (2006) that it should aim for "the best we can do by way of evidence-based policy." In its most traditional conceptualization, evidence-based policy refers to the development of policies informed by the most "robust and credible evidence" (Gertler et al., 2011).

decision-making related to its progress. Process evaluation refers to the assessment of program implementation and operation to inform its development and determine whether it unfolds as originally designed. Impact evaluation focuses on assessing program outcomes, by analyzing the program's impact in relation to its goals, and hence is interested in cause-and-effect relationships (cf. Chen, 1996; Edler et al., 2012). It can be used both for accountability purposes, that is, to justify public spending, or for improving program response (Stern et al., 2012). According to Rolfe (2016), the use of impact evaluation varies depending on whose interest the evaluation serves: policymakers might be more concerned about accountability and summative purposes, whereas practitioners might be more interested in improving the program response and formative evaluation. In this PhD thesis, I primarily focus on the latter. I return to this in Section 4.

As mentioned previously, evaluation is a contested field and multiple debates have emerged regarding the most appropriate approach to follow, leading to different perspectives in evaluation, also referred to as evaluation methodologies (Rolfe, 2016). Below, I discuss these perspectives, their epistemological basis, and different views of causality.

## 2.2 Exploring perspectives on evaluation[11]

Historically, scholars involved in epistemological discussion surrounding evaluation have drawn on three main philosophical paradigms:[12] post-positivism, which gained more attention during the 20th century and represents a shift from the positivism thinking; constructivism; and pragmatism (Alkin, 2013). This was also referred to as "paradigm wars" between positions along a positivism and constructivism spectrum, which seems to have slowed down and reached a "modest peace" in current days (Pawson, 2006). Accordingly, it seems that there is a general realization that different approaches in evaluation are needed "to tackle the evaluation of different types of programmes in their different stages of development" (Pawson, 2006, p. 14). In the following paragraphs, I discuss the different types of evaluation within this spectrum of positivism-constructivism approaches,

[11] Adapted from Haddad (2021b, pp. 25-29).

[12] According to Mathison (2005), a paradigm is "a worldview or perspective that, in the case of research and evaluation, includes conceptions of methodology, purposes, assumptions, and values… that typically consists of an ontology (the nature of reality), an epistemology (what is knowable and who can know it), and a methodology (how one can obtain knowledge)" (p. 289).

starting from the former approach, which still orients most of the methods-based evaluations (Khakee, 2003).

In a *positivistic* view (cf., e.g., Campbell, 1966; Cronbach and Shapiro, 1982), the systematic use of methods makes the study scientific (e.g., for determining accountability), and facts are preferred over value claims since they are perceived as rationally defined (Alkin, 2013). Underlying the positivistic view is the logic of experimentation and a "successionist" view of causality, which states that "the reason that we know one event caused another event is that the first event took place before the other event regularly" or, in short, "if p, then q; p, therefore q" (House, 2001, p. 311). This underlies the view of causality in experimental, for example, random control trials (RCT), and quasi-experimental designs, which look for the impact or causal effect of a program in outcomes, that is, the so-called counterfactual or, in other words, "what the outcome would have been for program participants if they had not participated in the program" (Gertler et al., 2011, p. 8).

According to Alkin (2013), adherents of *post-positivism* believe that truth can be approached but never reached, that is, one can measure the truth, but not uncover it, given that "a full understanding of truth can be approached but never reached" (p. 25). They also believe that reality can be studied objectively, as do positivists, but they differ in that they believe that reality cannot be understood in its totality (Christie and Fleischer, 2009). Their view of causality is that "causation is observable and that over time predictors can be established, but always some degree of doubt remains associated with the conclusion" (Christie and Fleischer, 2009, p. 24).

In *constructivism*, "claims, concerns and issues of stakeholders serve as organizational foci" (Guba and Lincoln, 1989, p. 50). These authors defend the idea of multiple realities and the view that stakeholders are involved in placing value. Guba and Lincoln (1989) call this constructivist approach the *Fourth Generation Evaluation*, which reflects the change in the role of the evaluator from an analyst of events, "re-constructor" of events, or "judge" to mediator and co-producer of social constructs (Amanatidou et al., 2014). This view follows an inductive logic, prefers qualitative methods, and sees causality as impossible to distinguish, given that the relationship between cause and effect is bidirectional, that is, everything affects everything else at once (Christie and Fleischer, 2009). This view of causality is also closely linked to the narrative

approach, which views participants as active agents in an intervention that can influence and generate successful outcomes (Stern et al., 2012).

*Pragmatists* point out that objectivity and subjectivity are "two positions on a continuum and argue that deductive and indictive logic should be used in concert" (Alkin, 2013, p. 17). Pragmatism is similar to constructivism in its understanding that "there are multiple explanations of reality, and at any given time there is one explanation that makes the most sense" (Christie and Fleischer, 2009, p. 25). However, it is related to post-positivism in its view of external reality and the lack of an absolute truth, as well as in its view of causality in the sense that it believes that causes may be linked to effects. However, pragmatists argue, absolute certainty about causation is impossible. Moreover, both quantitative and qualitative methods are legitimate (Christie and Fleischer, 2009).

Another philosophical position explored in evaluation, but not actually rooted in evaluation research, is *critical realism* (Pawson, 2013). Critical realism was developed by Roy Bhaskar, an English philosopher, and implies that the scientist's view is only a way to know reality and that observations are "not theory-neutral (it is influenced by our socio-political situatedness)" (Bryman and Bell, 2015, p. 61). According to Pawson (2013, p. 15), critical realism explains "nature's uniformities by unearthing the underlying mechanisms that give rise to them. To achieve this requires theory and it is the generative theories that allow us to both know how to manipulate the experiment and explain the results we then observe."[13] This philosophical position influenced, for instance, realist evaluation. Realist evaluation "is a species of theory-driven evaluation that holds the view that programs are theories incarnate" (Astbury, 2013, p. 385). Realist evaluation is considered the "European version" of theory-based[14] evaluation, primarily pushed forward by the Aspen Institute in the United States with the Theories of Change (ToC) framework (Connell et al., 1995; Fulbright-Anderson et al., 1998). However, while the realist evaluation is specified in realist terms and can be traced to the (European) critical realist movement, theories of change (and other

[13] In short, it derives from the notion that "to infer a causal outcome (O) between two events (X and Y), one needs to understand the underlying mechanism (M) that connects them and the context (C) in which the relationship occurs" (Pawson et al., 2005, pp. 21-22).

[14] Theory-based is also referred to in the literature as theory-led, theory-oriented, or theory-driven evaluation. Note that theory here does not refer to theory of science, but rather to the overall program theory, which explains how the program is supposed to work and how causal mechanisms will generate outcomes, which can either be based on stakeholders' views or be influenced by evaluators (Blamey and Mackenzie, 2007).

theory-based approaches) are in general silent about their epistemological and ontological views (Astbury, 2013).

Among the different evaluation methodologies advanced in the literature, as will be discussed further in the research methodology section (Section 4), the theory-based approach seems the most relevant among the multiple perspectives introduced above, because theory-based evaluation can potentially deal with complex and long chains of causality and explain how and why programs work rather than treating them as a "black box" (Rolfe, 2016). I further elaborate on theory-based evaluation below.

## 2.3 Theory-based evaluation[15]

Theory-based evaluation criticizes methods-driven and experimental-focused evaluations for lacking an explanation regarding how and why outcomes come about (Blamey and Mackenzie, 2007; Rolfe, 2019). It is "seen as providing a key to unlock complex processes between policy intent and policy outcome, by examining implementation, the causal processes that generate outcomes and contextual factors that influence them" (Rolfe, 2019, p. 294).[16] Rather than paying attention to methods, theory-based evaluations look at the processes, context, and outcomes that explain how a program is supposed to work. According to Stame (2004, p. 60):

> Theory-oriented evaluations present themselves as a new wave vis-à-vis method-oriented evaluations. In this new wave, what changes is the attitude toward methods. There are no more paradigm wars that are immobilizing the field; nor contention about pre-planned multi-method evaluations. All methods can have merit when one puts the theories that can explain a programme at the centre of the evaluation design. No method is seen as the "gold standard." Theories should be made explicit, and the evaluation steps should be built around them: by elaborating on assumptions; revealing causal chains; and engaging all concerned parties in this exercise.

To avoid confusion regarding the term "theory," which is used in a variety of ways in the theory-based literature (Blamey and Mackenzie, 2007), it is important to make some distinctions. In general, two main types of theory underlie policy evaluation: *program theory*, which refers to program implementation and to the mechanisms leading to outcomes detailing how programs work, and *social science theory*, which is a theory in the scholarly sense (Knill and Tosun, 2012).

[15] Adapted from Haddad (2021b, pp. 28-29/33-34).
[16] See also Chen (1990) and Weiss (1997).

Specifically, "program theory focuses on the nature of the evaluand itself (i.e., the program, treatment, intervention, policy, etc. being evaluated)" (Donaldson and Lipsey, 2006, p. 66). It relies on the assumptions that guide how the evaluation is implemented and will produce change (Astbury and Leeuw, 2010; Donaldson and Lipsey, 2006). Social science theory, in contrast, is not related directly to methods and practices of evaluation, but rather comprises scholarly theories that address the social phenomena underlying a social program; hence, it can be very relevant to evaluation (Donaldson and Lipsey, 2006; Knill and Tosun, 2012). This is also known as the *policy theory* underlying an intervention (cf. Molas-Gallart and Davies, 2006). I further elaborate on social science theory in Section 3.

Regarding the development of program theory, the most influential theory-based evaluation approaches are "Theories of Change" (ToC) and "Realistic Evaluation" (Blamey and Mackenzie, 2007).[17] The ToC approach aims at developing a visual and narrative model of the intervention under evaluation by analyzing how inputs and activities create expected outputs and how the intervention then generates long-term outcomes (Rolfe, 2019). According to Weiss (1995), a program's theories of change are a combination of implementation theory and program theory.[18] The former refers to how the program is carried out, that is, it tests the theoretical assumption of whether "the program is conducted as planned, with sufficient quality, intensity, and fidelity to plan" (Rogers and Weiss, 2007, p. 72). The latter in turn is related to the participants' responses to program activities (Weiss, 1997) and to the "hypothesized causal links between mechanisms released by an intervention and their anticipated outcomes" (Blamey and Mackenzie, 2007, p. 445).

Realist evaluation in turn is positioned "as a third, 'realist' way opposed to both the positivist-experimentalist and the nominalist-constructivist traditions in evaluation" (Alkin, 2013, p. 361). Realist evaluation can be both summative and formative, favors a multiple methods approach without any predefined preference for quantitative versus qualitative methods, and allows for the evaluation of both processes and impacts (Pawson and Tilley, 1997). While policy theory would

[17] In the literature, "Realist Evaluation" and "Realistic Evaluation" are used simultaneously to refer to the same evaluation approach.

[18] The literature uses different terms to describe the same type of theory or, concomitantly, similar terms to refer to theories that are different epistemologically (cf. Astbury and Leeuw, 2010; Blamey and Mackenzie, 2007; Knill and Tosun, 2012). Pawson and Tilley (1997), for instance, use the term "middle-range" theory to refer to Weiss' "program theory."

be similar to both ToC and realist evaluation, program theory in realist evaluation is mechanism focused (rather than process focused in ToC, which tries to examine the entire causal chain of program implementation) (Stern et al., 2012). Thus, program theory in realist evaluation is developed in realist terminology, which hypothesizes the mechanisms and triggers of this behavioral change, or, in the realist jargon: context, mechanism, and outcome configurations (CMOc). According to Pawson (2006, p. 25), these are the main sources of evidence, in which "interventions offer resources which trigger choice mechanism (M), which are taken up selectively according to the characteristics and circumstances of subjects (C), resulting in a varied pattern of impact (O)."

Regarding theory-based approaches, Rolfe (2019) states that the ToC approach does not follow a specific concept of causality, whereas realist evaluation follows a generative model of causation (Pawson, 2006; Pawson and Tilley, 1997). More specifically, when it comes to the assumptions surrounding causal attribution, the ToC approach follows the idea that, if the activities and outcomes predefined in the plan and agreed upon among different stakeholders as being plausible are delivered accordingly, then the attribution claim can be strengthened (Mackenzie and Blamey, 2005). In this way, Rogers and Weiss (2007) argue that "if the evaluation can show the series of micro-steps that lead from inputs to outcomes, then causal attribution for all practical purposes seems to be within reach" (p. 70). However, as they point out, it does not eliminate all threats to validity. Alternatively, Mackenzie and Blamey (2005) argue that another way to attribute outcome changes to the intervention is to combine the information gathered from the ToC with primary and secondary outcome data to determine whether the resulting change is due to the exposure of the target groups to the intervention. In this case, the data analysis can provide further evidence to attribute the outcomes to the intervention. Either way, "for both approaches it is important that the changes from the initial to the final theory articulation process are captured so that explanations are based on what has actually been delivered" (Mackenzie and Blamey, 2005, p. 163).

The RE approach in turn focuses on the generative causality perspective, which contrast with the successionist approach pushed forward by experimentalists, as discussed previously. Generative causality proposes explaining regularity generatively, in which, rather than trying to explain how variables or correlates are associated with each other, it explains how this association comes about (Pawson and Tilley, 1997). The generative mechanisms thus constitute the regularity itself.

Accordingly, realist evaluation "focuses on building and verifying a theory about how processes and mechanisms work in particular contexts to generate effects and changes" (Gates and Dyson, 2017, p. 31). Figure 2 shows the realist causal proposition, that is, the axiom that "causal outcomes follow from mechanisms acting in contexts" (Pawson and Tilley, 1997, p. 58).

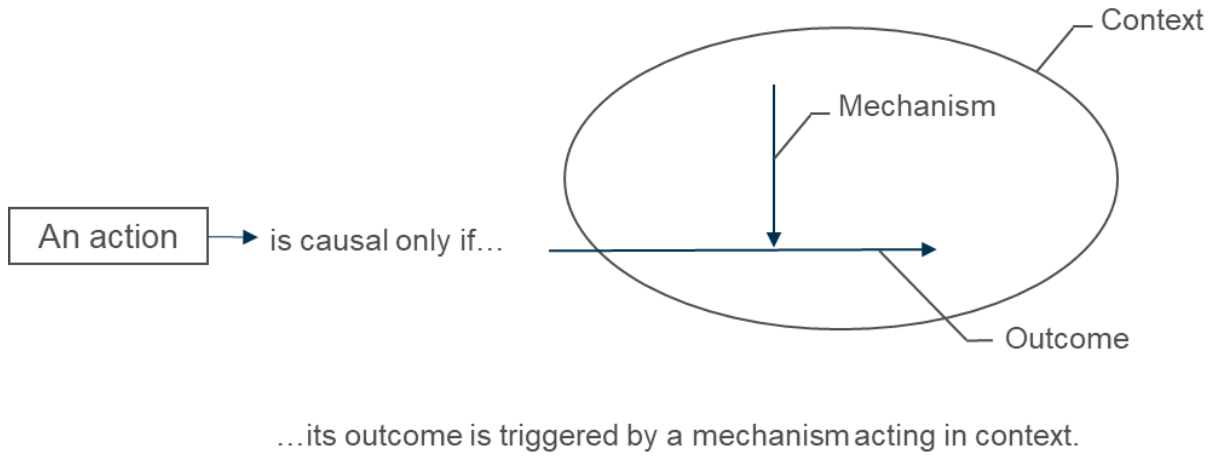


Figure 2. Generative causation. Source: Pawson and Tilley (1997, p. 58).

Some authors have also explored the possibility and feasibility of combining ToC and RE (Blamey and Mackenzie, 2007; Rolfe, 2016, 2019). According to Blamey and Mackenzie (2007, p. 451), ToC and RE could coexist in an evaluation, "with the former providing broad strategic learning about implementation and the latter bearing down on smaller and more promising elements of embedded program theory." Thereby, ToC emphasize the development of a complete theoretical model of interventions, whereas RE helps analyzing the causal mechanisms contributing to change (Rolfe, 2019).

## 2.4 Section summary

In this section, I defined public policy and evaluation and discussed different types of evaluation as well as the main evaluation methodologies described in the evaluation literature. Among these methodologies, the most relevant to this thesis, for reasons discussed further in Section 4, is the theory-based approach, which criticizes traditional evaluations focusing on methods and can potentially highlight how to build a program theory that accounts for the complexity underlying socio-technical system change. I then introduced the two most influential theory-based evaluation approaches, namely ToC and realist evaluation. I also pointed out that while ToC yields a broader

overview of the program constituents, realist evaluation sheds light on the mechanisms contributing to or hindering the generation of outcomes, and noted that some authors have also shown the potential of integrating the two.

In the next section, I will discuss the social science theory that can support the evaluation of TIP to complement and inform program theory. While this has been partially addressed in Section 1.1.1, where I discussed earlier framings of STI policy on which TIP builds, the types of change the TIP literature refers to relate to socio-technical system change, as conceptualized in the sustainability transitions literature. Therefore, in Section 3, I introduce the main approaches used to explain socio-technical transitions and account for recent efforts to address directionality and additionality.

# 3 Transformative innovation policy[19]

Following the review of policy evaluation, in this section, I discuss four main approaches that form the theoretical framing of sustainability transitions (Markard et al., 2012) to complement the discussion in Section 1, which covered the background related to STI policy in more detail. Together, both STI and the sustainability transition literature make up what the evaluation literature calls the social science theory, which is not specifically concerned with policy evaluation, but informs the type of outcomes and impacts one can look for when evaluating TIP. Later in the section, I discuss the main approaches put forward when analyzing directionality and additionality, which are the two main gaps this thesis aims to address when evaluating the effects of policies targeting transformative change.

## 3.1 Delineating transformative innovation policy

In Section 2.1, I discussed the definition of public policies and evaluation. I also stated that public policies are decisions made by governmental actors about taking actions or non-actions regarding a problem. In the of context innovation, the problem is usually related to a perceived lack of innovation—either in general or in terms of specific types of innovation (Weber and Rohracher, 2012; Woolthuis et al., 2005)—caused by market failures, systemic weaknesses, or transformational failures, depending upon which innovation policy framing one adheres to (see Section 1.1.1). Innovation policy in particular is a field within public policy that "comprises all combined actions that are undertaken by public organizations that influence innovation processes" (Borrás and Edquist, 2013, p. 1513).

The meaning given to innovation policy depends on what is meant by innovation. Innovation can be defined in a narrow sense, in which it corresponds to just an invention, or it could be defined in "a broader, more holistic perspective, which emphasizes the importance of looking at the entire innovation cycle from the creation of novel ideas to their implementation and diffusion" (Edler and Fagerberg, 2017, p. 4). This PhD thesis uses the term in the broader sense. According to these authors, the different perspectives on innovation have also been reflected in policy (and relate to the different framings of innovation policy as introduced in Section 1.1.1). For example, in the

[19] Partially adapted from Haddad (2021b, pp. 25-29).

1960s, the focus was on science, then it changed to technology and later to innovation, but it has also gone under the labels of industrial and research policy (Fagerberg, 2017).

The emergence of TIP as a third framing then builds on these previous conceptualizations of STI policy and targets the transformation of socio-technical systems, which might imply "different policies to those shaped by the traditional technological agenda" (Steward, 2012, p. 341). As already highlighted in Section 1.1.1, this means a shift from an economic to a societal policy agenda and from a narrow toward a broader understanding of the innovation process in comparison with previous STI policy framings (Diercks et al., 2019). Many authors emphasize that a policy mix for TIP should involve both demand- and supply-side instruments (Diercks et al., 2019; Rogge and Reichardt, 2016; Schot and Steinmueller, 2018), and support not only niche development but also the destabilization of incumbent regimes (Kivimaa and Kern, 2016).

As already discussed in Section 2.1, given that I use "policy" as an umbrella term, it refers to different levels of aggregation, including policy mixes, policy initiatives, programs, and projects. However, as I turn to evaluation, I primarily focus on the level of programs, which is a type of policy intervention "with an extensive range and scope that have activities that cut across sectors, themes, and geographic areas" (Leeuw and Vaessen, 2009, p. 10). Programs can be seen as composed of a portfolio of projects, which are single initiatives with specific objectives, and can also be grouped as a portfolio of programs (Arnold, 2004). While programs are not entire policy mixes as defined by Rogge and Reichardt (2016), they are composed of different elements of the policy mix, that is, policy strategy and instrument mix. This is not only because programs are usually the subject of evaluation rather than entire policy mixes, but also because of the empirical context that will be my object of study, as introduced in Section 4. Additionally, to date, most innovation policy interventions are still organized in the form of programs or portfolio of programs (Molas-Gallart et al., 2021).

## 3.2 Sustainability transition studies[20]

According to Zolfagharian et al. (2019, p. 2), "transition research is an interdisciplinary field, firmly rooted in the tradition of system thinking," centered around the socio-technical system as

[20] Adapted from Haddad (2021b, pp. 48-54).

the main object of research. Socio-technical systems refer to sectors such as transportation, water supply, and energy supply, which contribute to a specific societal function (Markard et al., 2012). If these systems undergo a shift in a set of transformative processes toward more sustainable modes of production and consumption, they are referred to as sustainability transitions. Transition studies also build upon a broad range of other theoretical approaches, including evolutionary economic theory (Nelson and Winter, 1985), social construction of technology (Bijker et al., 1984), and eco-innovation (Kemp, 2010). In broader terms, sustainability transitions are used to "refer to large-scale societal changes, deemed necessary to solve grand societal challenges" (Loorbach et al., 2017, p. 600).

Below, I discuss four of the dominant schools of thought related to socio-technical transitions: strategic niche management (SNM), transition management (TM), the multi-level perspective (MLP), and technological innovation systems (TIS) (Loorbach et al., 2017; Markard et al., 2012). Together with previous framings of STI policy, discussed in Section 1.1.1, the literature on sustainability transitions complements what would compose the social science theory that underlies TIP evaluation.

### 3.2.1 Strategic niche management

*Strategic niche management* was introduced at the end of the 1990s and assumes that "sustainable innovation journeys can be facilitated by modulating of technological niches, that is, protected spaces that allow nurturing and experimentation with the co-evolution of technology, user practices, and regulatory structures" (Schot and Geels, 2008, p. 538). SNM involves a clear governance aspect in that it suggests that strategically managing niches is "a possible (or even necessary) strategy for governments to manage the transition process to a different regime" (Kemp et al., 1998, p. 185). It was suggested early on in transition studies as a way to enable regime shifts (Markard et al., 2012). Regimes in turn were called "technological regimes" in earlier work, referring to a set of engineering practices, process technologies, skills, product characteristics, etc., embedded in broader institutions and infrastructures (Rip and Kemp, 1998).[21] These technological

[21] This in turn widens the concepts of "technological regimes" from Nelson and Winter (1982), which indicated that cognitive routines are embedded in the minds of engineers (Geels, 2002).

regimes support and impose a direction of incremental change along established trajectories (Geels, 2002).

A general argument in SNM is that protected spaces are required for entrepreneurs and system builders to experiment with new technology in relation to user practices, demonstrate its viability, and attract funding, as well as to achieve the institutional adaptations needed to eventually achieve a widespread diffusion (Schot and Geels, 2008). There are several conceptualizations of niche development, including the early work by Kemp et al. (1998), as well as later elaborations of their framework by other scholars (e.g. Geels and Raven, 2006; Schot and Geels, 2008; Smith et al., 2010), which identify three main niche development processes: learning processes, articulation of expectations and visions, and the enrollment of commitments from a growing network of actors.[22] In more recent literature, three properties of niches as protected spaces have been identified (Smith and Raven, 2012). First, shielding implies that niches protect emerging innovation from selection pressures in the mainstream market or other relevant selection environments (Smith and Raven, 2012) and thus create a space for experimentation (Verhees et al., 2013). Second, nurturing corresponds to the three main niche development processes described above (Naber et al., 2017; Raven et al., 2016; Verhees et al., 2013). Third, empowering refers to different processes that improve the competitiveness of niche innovations and remove shielding, either by adapting the niche innovation to fit current selection environments (fit-and-conform processes) or institutionalizing shielding by making mainstream selection environments more agreeable to the niche innovation (stretch-and-transform processes) (Raven et al., 2016; Verhees et al., 2013).

### 3.2.2 Transitions management

*Transitions management* focuses on long-term sustainable transformation processes and emphasizes the notion of reflexive governance, that is, "knowledge integration, anticipation of long-term systemic effects, adaptivity of strategies and institutions, iterative participatory goal formulation and interactive strategy development" (Kemp and Loorbach, 2006, p. 103). TM is a model for governance and management strategy based on a set of general principles, including the

[22] Kemp et al. (1998) identified three aims of strategic niche management: (i) to articulate necessary technological and institutional changes and adaptations, (ii) to set learning processes in motion in relation to different technological options, (iii) to stimulate the development and diffusion of these and other complementary technologies, and (iv) to build a semi-coordinated constituency around new technology.

need to deal with uncertainties, keep options open and manage fragmented policies, have a long-term orientation, take into consideration the right scale (i.e., international, national, local) when finding solutions, and set specific tasks for the government (Loorbach and Rotmans, 2010; Rotmans et al., 2001). These principles are translated into an operational model, namely the transition management cycle, which encompasses (i) structuring the problem and the transition arena, (ii) developing a transition agenda and deriving the transition paths, (iii) carrying out experiments and mobilizing actors, and (iv) monitoring, evaluation, and learning (Loorbach and Rotmans, 2010).

### 3.2.3 Multi-level perspective

The *multi-level perspective* builds upon concepts such as SNM (Hoogma, 2000; Kemp et al., 1998; Smith, 2003) and TM (Kemp and Loorbach, 2006; Kemp and Rotmans, 2005). In the MLP framework, transitions are conceptualized as major changes in the socio-technical configurations through which important sectoral societal functions are fulfilled (Geels, 2002, 2004), which unfold at multiple levels: niche, regime, and landscape (Geels, 2002). Socio-technical transitions are dependent on the development and upscaling of new technologies and solutions. In the transition literature, this is assumed to happen through the gradual build-up and institutionalization of socio-technical *niches*. Niches can be thought of as "protected spaces," which temporarily shelter emerging innovations from mainstream selection pressures (Smith and Raven, 2012; Smith et al., 2010). They allow promising technologies to be developed and used in an experimental setting, where technology, user practices, and regulations can be explored in a co-evolutionary way (Schot and Geels, 2008). The *socio-technical regime* builds upon the "technological regime," as mentioned above, and understands that rules are embedded in a wider knowledge base, structures, processes, and practices, influenced by engineers, as well as other social groups (Geels, 2002). Socio-technical regimes are characterized by the stability and inertia of established socio-technical configurations, which originate from socio-technical systems, actor networks, and regime rules (Geels, 2004). *Landscape* in turn refers to the heterogeneous and external factors affecting both niches and regimes, such as oil prices, wars, emigration, and environmental problems (Geels, 2002). In general, it can be defined as "the set of residual factors that have an impact on innovation and production processes without being influenced by the outcome of innovation processes on a short and mid-term basis" (Markard and Truffer, 2008, p. 606).

In the MLP, transitions occur due to the interplay of processes in these three multi-levels. Transitions can unfold in many ways, as they can happen gradually, on a step-by-step basis, or chaotically, involving the disruption of the existing regime and the replacement of key technologies (Lindberg et al., 2019). These many ways can be analyzed in terms of *transition pathways* (Geels et al., 2016; Geels and Schot, 2007). Geels and Schot (2007) distinguished four transition pathways in addition to the "zero proposition," wherein the regime remains stable and reproduces itself: (0) *reproduction process* represents the business-as-usual process, in which the regime is reinforced and stabilized and no niche innovations break through, even if they are present; (1) *transformation path* refers to moderate landscape pressures that happen when niche innovations have not yet been developed, leading to a reorientation of the regime by regime actors; (2) *de-alignment and re-alignment* develops when there is a major landscape pressure that erodes the regime (i.e., de-aligns it) and allows for multiple niche innovations to compete; when one niche innovation becomes dominant, it takes over the regime and re-aligns it; (3) *technological substitution* refers to major landscape pressures that occur when disruptive niche innovations are sufficiently developed to substitute the existing regime; and (4) *reconfiguration* refers to the adoption of symbiotic niche innovations by the regime, which is further adjusted in the event of landscape pressures. A fifth proposition accounts for shifts between pathways, in which disruptive landscape pressures lead to a sequence of transition pathways, for example, starting from transformation and then shifting to substitution or de-alignment and re-alignment (Geels and Schot, 2007). This transition pathways typology was later reformulated to reflect the reorientation of the socio-technical regime as being a result of "landscape pressure, societal debates and tightening institutions," as well as of "the moves and countermoves of actors and social groups, which are constrained by 'rules of the game' and oriented towards reproducing or modifying elements of socio-technical systems" (Geels et al., 2016, p. 897).

### 3.2.4 Technological innovation system

Together with the MLP, the *technological innovation system* framework is another strand of literature that analyses the radical innovation leading to transformation processes, also rooted in

evolutionary theory[23] (Markard and Truffer, 2008). While initially the TIS concept had little to do with sustainability and focused on providing policymakers with tools to promote specific technologies targeting economic growth (cf. Andersson, 2020; Carlsson, 1995), the concept was also applied to investigate the emergence and growth of renewable energy systems (cf. Johnson and Jacobsson, 2001). Thus, TIS "emerged as an influential concept in academic debates on the design of policies to stimulate environmental innovations and facilitate sustainability transition" (Magnusson and Berggren, 2018, p. 217).

A technological (innovation) system can be defined as "a network of agents interacting in a specific *economic/industrial area* under a particular *institutional infrastructure* or set of infrastructures and involved in the generation, diffusion, and utilization of technology" (Carlsson and Stankiewicz, 1991, p. 111). This implies that TISs are problem-solving knowledge networks (rather than production networks) related to particular product markets (Bergek, 2019). Additionally, a TIS comprises all the elements that influence the innovation system, and not only the focal technology. Moreover, the TIS may also be part of a sub-system of a sectoral system or cut across different sectors, and may have a geographical dimension (Bergek et al., 2008).

In the TIS literature, innovation outcomes have been conceptualized in both structural and functional terms. Some literature describes processes that contribute to the structural build-up of new systems, such as actor entry, network formation, and institutional adaptation (Jacobsson and Bergek, 2004; Jacobsson and Johnson, 2000).[24] Regarding functionality, seven key processes have been identified that contribute to the development, diffusion, and utilization of new technologies and thus to changes in the socio-technical system of a sector (as mentioned in Section 1.1): knowledge development and diffusion, entrepreneurial experimentation, the guidance of the direction of search, market formation, legitimation, resource mobilization, and development of positive externalities (Bergek et al., 2008), detailed in Table 1. To analyze a TIS and help policymakers in the selection and prioritization of public policies, Bergek et al. (2008) proposed an analytical scheme composed of six main steps: (i) defining the TIS in focus; (ii) identifying the

[23] This is because both approaches highlight aspects such as the "importance of networks and learning processes together with the crucial role of institutions in successful innovation processes. Both acknowledge phenomena such as path-dependency, lock-in, interdependence, non-linearity and coupled dynamics" (Markard and Truffer, 2008, p. 597).

[24] Some authors also include the accumulation of knowledge and artefacts among the structural processes (cf., e.g. Bergek et al., 2008)

structural components of the TIS; (iii) mapping the functional pattern of the TIS; (iv) assessing the functionality of the TIS and setting process goals; (v) identifying inducing and blocking mechanisms; and (vi) specifying key policy issues.

Table 1. Technological innovation system functions and typical indicators for assessment.

| Function | Description[a] | Hekkert et al. (2007) | Bergek et al. (2008) and Bergek (2019) |
|---|---|---|---|
| **Knowledge development**<br><br>**Knowledge diffusion** | Broadening and deepening of the knowledge base of a TIS, sharing of knowledge between actors within the system and new combinations of knowledge because of these processes. | R&D projects over time<br>Patents<br>Investments in R&D<br>Number of workshops and conferences on a particular technology<br>Network size and intensity over time | Bibliometric indicators (citations, volume of publications, orientation).<br>Learning curves<br>Subsidies for demonstrations<br>Technological collaboration, communication, and dissemination<br>Learning by doing and using<br>Knowledge acquisition, absorption, and accumulation<br>Learning by doing and practicing |
| **Entrepreneurial experimentation** | Problem-solving and uncertainty reduction through real-world trial-and-error experiments at different scales with new technologies, applications, and strategies. | Number of new entrants<br>Number of diversification activities<br>Number of new experiments with a new technology | The breadth of technologies used, and the nature of the complementary technologies employed<br>Number of incubators<br>Advice systems for SMEs<br>Pilot and demonstration with new technologies |
| **Market formation** | The opening up of a space or an arena in which goods and services can be exchanged in (semi-)structured ways between suppliers and buyers, including, for example, articulation of demand and preferences, product positioning, standard-setting, and development of rules of exchange. | Number of niche markets introduced<br>Specific tax regimes for new technologies<br>New environmental standards that improve the chances for new environmental technologies | Who are the users and what they are purchasing?<br>Demand articulation<br>Market size and customer group<br>Actors' strategies, role of standards and purchasing processes |

| Function | Description[a] | Hekkert et al. (2007) | Bergek et al. (2008) and Bergek (2019) |
|---|---|---|---|
| **Influence on the direction of search** | Mechanisms that influence to what opportunities, problems and solutions firms and other actors apply their resources, incentivizing and pressuring them to engage in innovative work within a particular technological field and determining what strategic choices they make within that field. | Specific targets set by governments or industries regarding the use of a specific technology<br>Number of articles in professional journals that raise expectations about new technological development | Beliefs in growth potential<br>Incentives from factor/product prices<br>The extent of the regulatory pressures<br>The articulation of interest by leading consumers |
| **Resource mobilization** | The system's acquisition of different types of resources for the development, diffusion and utilization of new technologies, products, and processes, most notably capital, competence and manpower, and complementary assets (e.g., infrastructure). | Funds made available for long-term R&D programs set up by industry or government to develop specific technological knowledge<br>Funds made available to allow testing of new technologies in niche experiments<br>Perception of the actors regarding the access to sufficient resources | Rising volume of capital, increasing volume of seed and venture capital (financial resources)<br>Changing volume and quality of human resources<br>Changes in complementary assets and infrastructure<br>Public and private funding<br>Education and training<br>Development of complementary infrastructure |
| **Legitimation** | The process of gaining regulative, normative, and cognitive legitimacy for the new technology, its proponents, and the TIS as such in the eyes of relevant stakeholders, that is, increasingly being perceived as complying with rules and regulations, societal norms and values, and cognitive frames. | Rise and growth of interest groups<br>Lobby actions | Alignment between TIS and the current legislation<br>How legitimacy influences demand, legislation, and firm behavior<br>Technology validation and standardization<br>Changing norms and values |
| **Development of positive externalities** | The creation of system-level utilities (or resources), such as pooled labor markets, complementary technologies, and specialized suppliers, which are also available to system actors that did not contribute to building them up. | - | Emergence of pooled labor markets<br>Emergence of specialized intermediate goods and service providers<br>Information flows and knowledge spillovers<br>Development of complementary technologies |

[a] Elaboration on Bergek (2019) and Bergek et al. (2021), which draw on Bergek et al. (2008).

## 3.3 Directionality and additionality

In this section, I focus on directionality and additionality. In Section 1.2, I highlighted these compose two main gaps related to capturing the effects of policy targeting transformative change. Below, I describe the few approaches that were put forward with respect to TIP.

Regarding directionality, Section 1.2 highlighted that some of the TIP evaluation frameworks introduced in Section 1.1.2 address directionality for policy analysis but do not discuss explicitly how to assess it. In addition, other approaches targeting directionality have been proposed in the sustainability transitions literature that were not specifically developed for evaluation purposes, but which can also shed some light on how directionality can be influenced and understood. One approach is the transition pathways proposed by Geels and Schot (2007) and Geels et al. (2016), as already discussed in Section 3.2.3. Another approach has been put forward by Yap and Truffer (2019, p. 1030), who propose an analytical framework based on the TIS approach to better understand "whether, how and by whom the directionality of innovation systems can be influenced." While directionality is related to the interplay of different functions, the authors focus on analyzing the guidance of the search function[25] and how actors influence different layers of the selection environment. Additionally, Pel et al. (2020) analyze how to govern "transitions directionality," that is, the different possible socio-technical change paths, by combining insights from socio-technical multiplicity, divergent normative appraisals, and process dynamics.[26] These three angles are used together to trace challenges and enable a "directionality-conscious transitions governance."

An alternative framework would be to specify the "problem directionality" and its corresponding "solution directionality," as proposed by Wanzenböck et al. (2020) for mission-oriented innovation policies. The authors propose this typology to decompose societal challenges based on their level of contestation, complexity, and uncertainty. Accordingly, the authors argue that problem-solutions can either converge or diverge and propose three different pathways in moving toward

[25] Yap and Truffer (2019, p. 1031) argue that the guidance of the search function has a more explicit connection with directionality and represents the "top-down and bottom-up activities that different actors entertain in order to shape the sectoral selection environment in favour or against alternative trajectories."

[26] According to Pel et al. (2020, p. 4): (i) socio-technical multiplicity refers to the socio-technical configurations pursued by actors; (ii) appraisal diversity relates to the "full range of appraisals and evaluation schemes through which configurations are valued and compared"; and (iii) process dynamics is the experienced or anticipated turns toward the development trajectory.

converging approaches to problem-solution combinations. Andersson et al. (2021) propose a "morphology" of socio-technical systems to understand the outcomes of directionality. This is composed of three main dimensions related to the boundaries of the investigation: (i) temporal, which can be related to prospective or retrospective studies; (ii) spatial, that is, studies can focus on the local, national, or global level; and (iii) socio-technical, which is multidimensional and includes the technical, social, and spatial (or geographical space) configurations.

Other authors have also considered directionality in the design and implementation of TIP (Bergek et al., 2023; Cevallos and Merino Moreno, 2022; Parks, 2022), or have sought to assess the directionality of socio-technical pathways rather than policy (Geels et al., 2016; Turnheim et al., 2015), or even to calculate an index to measure the directionality of regions in tackling societal challenges (Cappellano et al., 2022). Although these frameworks and approaches have yielded interesting insights into the conceptualization of directionality, discussion of how to address it in evaluation is ongoing. Apart from Andersson et al. (2021), which touches upon the outcomes of directionality, how evaluators can assess this concept in policy evaluations remains understudied.

Concerning additionality, Section 1.2 pointed out that the concept has hardly been adopted in the TIP literature, with a few exceptions related to addressing behavioral additionality in the context of grand challenges (Amanatidou et al., 2014) and capturing additionality at the system level (Janssen, 2019). In this section, I discuss two main pieces of work that, although not directly related to additionality, touch upon causality and explanation in sustainability transitions approaches and, hence, can shed light on the additionality issue. Both papers refer to the identification of mechanisms in sustainability transition frameworks, which also connects with the theory-based approach to evaluation.

The first is de Oliveira et al. (2020), who refine the TIS framework using a mechanism-based approach to improve its explanatory power concerning systemic malfunctioning and its implications. According to the authors, this would imply a clearer description of causes, contextual conditions, and outcomes. Causal mechanisms, or causal pathways between cause and outcomes, are represented by how systemic problems hinder the fulfilment of TIS functions, that is, blocking

mechanisms.[27] Outcomes are the patterns of system functioning, "given by the hindrance of system processes" (de Oliveira et al., 2020, p. 25). Contextual conditions in turn involve factors as indicated by (Bergek et al., 2015, p. 45), for example, TIS context structures (i.e., "all other structures and relevant factors outside the TIS"), including technological, sectoral, geographical, and political, as well as TIS-context interactions. While these mainly represent factors outside the TIS boundary, other contextual factors might also include TIS-endogenous conditions, that is, internal to the TIS boundaries (de Oliveira et al., 2020; de Oliveira and Negro, 2019). In other words, "blocking mechanisms are the 'pathways' caused by one or multiple systemic problems that yield an inadequate fulfilment of system processes under specific contextual conditions" (de Oliveira et al., 2020, p. 26).

The second is Geels (2022), who suggests that the sustainability transitions framework already uses some elements from critical realism and, in terms of causality, accommodates a notion of complex causality that can explain long-term transformative change. For example, the TIS functions have the analytical basis to make both formative and summative assessments of many of the mechanisms related to socio-technical transition processes (Markard and Truffer, 2008). The MLP and SNM in turn address niche and regime-level processes needed for a transition to unfold (Markard et al., 2015; Markard and Truffer, 2008; Smith and Raven, 2012), which the TIS framework does not necessarily address. Specifically, Geels (2022, p. 8) distinguishes between three kinds of complex causality:

> Conjunctural and event-chain causalities can explain the temporal unfolding of transitions. Configurational and event-chain causalities can help explain endogenous transformation processes. And configurational causality can address the heterogeneity of socio-technical systems.

## 3.4 Section summary

This section complemented the social science theory underlying TIP evaluation. I first delineated TIP and highlighted that in this thesis I mainly focus on TIP program evaluation. I then discussed

[27] This is similar to the definition of blocking mechanisms provided by the TIS literature (Bergek et al., 2008; Wieczorek and Hekkert, 2012), in which blocking mechanisms emerge as a result of endogenous negative attributes of structural elements, e.g., lack of actor capability. However, de Oliveira et al. (2020) propose that the concept needs to be refined, first, to include exogenous negative attributes, e.g., poor infrastructure or weak regulatory alignment. Second, it needs to better explain "how blocking mechanisms relate to systemic problems (causes), how they come up and manifest themselves and how they lead to poor system functioning" (de Oliveira et al., 2020, p. 25).

the main approaches used in the sustainability transitions literature, which complements the literature on previous framings of STI policy described in Section 1.1.1. Specifically, I introduced four main approaches: TM, SNM, MLP, and TIS. This literature aims to explain transformative processes that lead to shifts in entire socio-technical systems, such as the transport, energy, and water supply sectors. These processes thus inform the type of process evaluators should look for when performing evaluation of TIP and can thus shed light on the type of outcomes evaluators should be considering. This connects with the concepts of directionality and additionality, which are also described in this section. In this regard, I introduce some of the recent approaches put forward toward addressing directionality and additionality in relation to TIP, which can potentially shed light on how to address these concepts in evaluation.

Together, the literature on policy evaluation (Section 2) and the social science theory (Section 1 and Section 3) informing it provide a starting point to understand the distinguishing characteristics of an "ideal type" TIP evaluation (RQ1), as well as to move toward addressing directionality and additionality (RQ2). In the next section, I discuss my research methodology while detailing my research process, choice of evaluation methodology, and methodological choices to address the research questions introduced in Section 1.2.

# 4 Research methodology

In Sections 2 and 3, I outlined the most relevant literature on program theory and social science theory. In this section, I describe my research methodology. First, I introduce the broader research project in which this PhD thesis is embedded. I then briefly describe my research process and specify my research design to address the research questions introduced in Section 1.2. Finally, I reflect on my main methodological limitations.

## 4.1 The broader research project[28]

This thesis is part of a research project funded by Vinnova, the Swedish Innovation Agency. Vinnova is the central player in terms of innovation policy in Sweden and has through the years been responsible for science-industry and inter-firm collaborations, as well as for the promotion of technology transfer activities (OECD, 2016). Vinnova has in the past years enforced both its policy discourse and resource allocation in "Grand Challenges" (Grillitsch et al., 2019). The agency has been gradually integrating the sustainable development goals (SDGs) of the 2030 Agenda adopted by the United Nations into its operations (Vinnova, 2021b). Therefore, the agency aims to build innovation capacity in Sweden, and contribute to sustainable growth. Its current vision is "that Sweden is an innovative force in a sustainable world" (Vinnova, 2021b).

Vinnova has been funding the Swedish Transformative Innovation Policy Platform (STIPP) since 2018, with the overall aim of advancing "the understanding of the dynamics and governance of sustainability transitions" (STIPP, 2019). The platform involves the participation of senior and junior researchers from four Swedish universities: Lund University, Jönköping International Business School, Chalmers University of Technology, and Linköping University. In 2021, the *Laboratoire Interdisciplinaire Sciences Innovations Sociétés* (LISIS), based in France, also joined STIPP. The platform focus encompasses two main transition areas of strategic importance to Sweden: the transformation toward a bio-based economy and the development of smart cities (STIPP, 2019). Overall, STIPP is organized around six interrelated research projects (RPs) that contribute to five different overarching work packages (WPs), as shown in Figure 3.

[28] Adapted from Haddad (2021b, pp. 21-23).

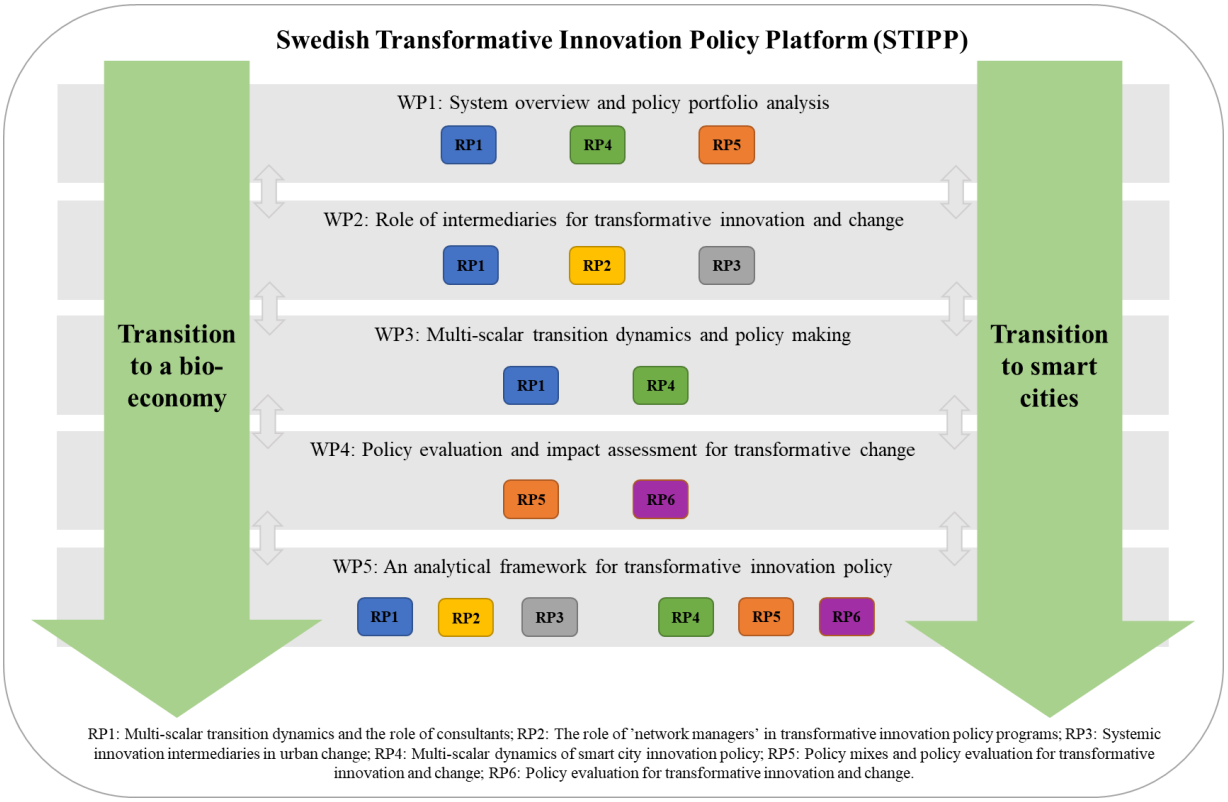


Figure 3. Swedish Strategic Innovation Program Platform STIPP organization and rationale.

More specifically, this thesis focuses on RP6: Policy evaluation for transformative innovation and change, a part of WP4: Policy evaluation and impact assessment for transformative change. The starting point of this project is the fact that current approaches to evaluating innovation policy are not sufficiently adapted to assess transformative change. The overall objective of RP6 is to study how transformative change can be assessed and evaluated considering the realities of both policy practitioners and network managers. Empirically, the original project intended to focus on specific policy initiatives, such as the Strategic Innovation programs (SIPs), [29] possibly in comparison with more traditional innovation policy initiatives in the same empirical field, such as the Challenge-Driven Innovation (CDI) and Vinnväxt programs.[30]

[29] The SIPs are coordinated by Vinnova, with the support of the Swedish Energy Agency and the Swedish Research Council Formas, together with participating organizations, and consist of 17 areas that are considered of strategic importance to Sweden (Vinnova, 2018).

[30] The Vinnväxt program focuses on "promoting sustainable regional growth by developing internationally competitive research and innovation milieus in specific growth fields" (Vinnova, 2016, p. 4). The Challenge-Driven Innovation program in turn promotes the development of sustainable solutions by providing funding for collaborative projects that aim at developing solutions to societal challenges and contributing to the SDGs (Vinnova, 2019).

## 4.2 The research process

First, I read the literature on TIP. Together with my supervisor, Anna Bergek, co-supervisor, Hans Hellsmark, and colleague, Valentina Nakic, we discussed the possibility of developing a literature review of the policymaking process of TIP. Additionally, Valentina and I were involved in a literature course organized by Anna and Hans that touched upon the earlier framings of STI policy, policy mixes, sustainability transitions, and TIP in general.

While both the literature review and the course gave me a point of departure in relation to TIP and also shed light on earlier efforts toward evaluating STI policies, it also showed that efforts toward TIP evaluation mostly relied on previous approaches to STI policy evaluation, which were already lagging advances in innovation policy theory. Additionally, most papers did not refer to the literature on policy evaluation and some performed a policy analysis rather than an impact evaluation in which causal-effect relationships should be accounted for. Nonetheless, these papers offered insights into the identification of policy effects and, specifically, on directionality and additionality. At this point, I realized that I needed to further read the literature on policy evaluation to look at different perspectives put forward in evaluation practice that could help address the issues of TIP impact evaluation. Therefore, I also engaged in reading the different perspectives on policy evaluation.

While trying to link back with many proponents of formative evaluation when addressing innovation policy evaluation, I opted to study the theory-based evaluation literature in greater depth, given that many evaluators would argue that this could help deal with the complex policy programs. Besides, theory-based evaluations are an alternative to traditional methods-oriented evaluation practices that underlie most of the additionality calculations in STI policy evaluation yet cannot explain how and why programs work. I further discuss the choice of an evaluation methodology in the next section.

From these readings, Anna and I sketched a first draft of a framework for TIP program evaluation that could capture directionality and additionality. In parallel and as part of the earlier drafts of this framework, we also discussed the type of processes evaluators should look for in the context of TIP. Departing from the TIS functions and reading the criticisms of the approach and the possibility of integrating it with insights from the MLP, we discussed the opportunity of adapting

the functions framework to cover additional transition processes not yet been emphasized within the framework. Both of us engaged in a process of identifying key innovation and transition processes that could better inform the types of transformative changes we could look for in TIP evaluation, while addressing directionality and additionality.

While these earlier efforts involved mostly conceptual and methodological work, which were presented in different conferences and workshops (see Section 5), research practice was also needed. Therefore, during my Licentiate writing in 2021, I took the first steps toward applying the framework for evaluating TIP programs. Given that the STIPP platform was set around the transition toward bioeconomy and smart cities, I departed from the BioInnovation SIP case. In this first attempt, I mostly used secondary data.

After my Licentiate, I took further steps to engage with policy practice. First, in May 2022, I engaged with Emily Wise and Erik Arnold from the Governing Resilience and Transformation (GReaTr) project, led by Lund University, in developing a series of case analyses for Vinnova. Within GReaTr, one of the foci of their work packages is the evaluation and design of transformative innovation policy through cross-case analysis in relation to the SIPs, CDI, and Vinnväxt programs. Given the synergy between this with my work package (WP4) from STIPP, the collaboration made sense and served as an alternative to comparing BioInnovation with other similar programs within CDI and Vinnväxt, as originally planned within the STIPP project. Rather, I could look at all the overall innovation evaluation practice developed by Vinnova in these three innovation policy programs to identify the main challenges in translating TIP evaluation theory into practice.

Second, further in May 2022, I engaged in an internship at the OECD in Paris. There, together with Michael Keenan, I became involved in S&T Policy 2025, which looks at how STI policy can enable sustainability transitions. Third, after my internship, I returned to Sweden and returned to examining the BioInnovation SIP. Initially, the idea was to apply the framework Anna and I had developed to BioInnovation. However, after my PhD Final Seminar held in April 2023 with Eugenia Perez Vico and based on her feedback, I decided to change the scope of this application to reflect on key methodological challenges that evaluators could encounter when performing a TIP program evaluation. Anna guided me during this process.

## 4.3 Research design

### 4.3.1 Identifying distinguishing characteristics of an "ideal type" TIP evaluation

To identify the distinguishing characteristics of an "ideal type" TIP evaluation and address RQ1, I combine insights from Papers I and II. Specifically, Paper I presents a systematic literature review of transformative innovation policy, which was developed to synthesize the academic work around TIP, based on preliminary research that had been done in trying to understand what TIP is in relation to previous innovation policy framings (Diercks et al., 2019; Schot and Steinmueller, 2018). The paper follows Petticrew and Roberts (2008) to conduct systematic literature reviews. According to these authors, "systematic reviews are literature reviews that adhere closely to a set of scientific methods that explicitly aim to limit systematic error (bias), mainly by attempting to identify, appraise and synthesize all relevant studies (of whatever design) to answer a particular question (or set of questions)" (p. 9).

We began the systematic literature review by defining keywords related to TIP, which included, terms related to transformative change, sustainability transitions, and grand challenges. Based on these keywords, we searched the Scopus database. After screening the resulting search and defining the eligibility and the criteria for inclusion of the papers, we ended up with 45 papers, plus an additional one via snowballing. By analyzing the 46 papers, we first identified a set of nine characteristics of TIP, which we aggregated into five thematic areas that cover distinguishing features of TIP in relation to previous STI policy framings. We also went through each paper to identify the mention of different stages of the policy cycle, which is one of the most used methods of explaining the policymaking process (Cairney, 2012). We used a generic policy cycle with six stages: agenda-setting, policy formulation, legitimation, implementation, monitoring and evaluation, and policy learning, adapted from Cairney (2012) and Howlett and Giest (2013). These latter two stages, that is, monitoring and evaluation and policy learning, are relevant for policy evaluation in general. As a result of this process, described in more detail in Paper I, we identified a set of challenges related to each stage of the policy cycle and how these were related to the five thematic areas. These challenges (namely "first-order" challenges) were then aggregated into "second-order" challenges, which would represent overarching TIP challenges that cut across different thematic areas and stages of the policy cycle. The identified thematic areas and first- and

second-order challenges related to the monitoring and evaluation and policy learning stages, then, offer the first insights around the distinguishing characteristics of TIP evaluation.

Paper II examines the practice of innovation policy evaluation in Europe, taking into consideration categories (grouped into key dimensions) that characterize evaluations (Edler et al., 2012) or attributes for a system-level evaluation (Borrás and Laatsit, 2019) (see Section 1.2). However, we found no study covering what would characterize TIP evaluation. To derive key characteristic of what would encompass an "ideal type" TIP evaluation, we then draw on the categories from Edler et al. (2012), who proposed five dimensions for analyzing the practice of innovation policy: evaluation set-up, policy measures, main questions, methods used in the evaluation, and policy and quality. According to the authors, these dimensions, further subdivided into different categories, serve as a first step toward a typology for innovation policy evaluation. Thus, we used these dimensions and categories as a starting point to derive those different for TIP evaluation, based on insights from the TIP literature. A more detailed account of the development of the analytical framework can be found in Paper II.

I then compared the TIP evaluation characteristics found in the literature review presented in Paper I and the framework of a TIP program evaluation strategy from Paper II to derive the distinguishing characteristics of TIP evaluation. Most notably, these papers highlight the need to address directionality and consider system-level processes, which are considered more in-depth in Papers III and IV.

### 4.3.2 Accounting for directionality and additionality

The second research question addressed in this PhD thesis relates to directionality and additionality and how they can be accounted for in TIP program evaluation. In Section 2, I pointed out that among the different perspectives on policy evaluation, theory-based approaches seemed the most suited to addressing the complexities underlying TIP program evaluation. In this section, I further motivate the selection of this approach and also describe the development of an integrated framework for evaluating TIP, as well as the identification of the transformative processes evaluators should look into when developing impact evaluations for TIP programs.

#### *4.3.2.1 Choosing an evaluation methodology*

As mentioned in Section 4.2, during the literature review I realized that the literature on evaluation was not often considered among the recent developments on STI policy evaluation. Therefore, to identify a way to address directionality and additionality, I went back to the policy evaluation literature to study the different evaluation perspectives put forward by evaluation practice. This was necessary because evaluation is a contested field underpinned by different opinions and philosophies of science, as discussed in Section 2.2, with many discussions surrounding the spectrum between positivism and constructivism and between quantitative versus qualitative methods (Rolfe, 2016).

In the innovation policy literature, the evaluation of additionality has historically focused on quantitative methods (Rohracher et al., 2022), using approaches like "bang for the buck" and other quasi-experimental econometric methods (Janssen, 2019). Quantitative methods also seem to be favored recently by a few proponents of TIP evaluation, who "emphasize that rigorous evaluations require quantitative analysis of outcomes with respect to a counterfactual" (Marques Santos and Coad, 2023, p. 7). However, when it comes to transformative innovation policy, a clear counterfactual is lacking due to the fact that TIP interventions aim at promoting systemic changes; thus, their effects might go beyond the unique set of beneficiaries of the intervention itself (Janssen, 2019). Therefore, the traditional counterfactual approach based on control groups was deemed unsuitable for understanding the outcomes of TIP. Other TIP scholars propose formative evaluation and the use of flexible ToCs, in which the evaluation is developed with the participation of stakeholders with the aim of improving the implementation of the intervention (Boni et al., 2019; Molas-Gallart et al., 2021), which resembles the constructivist perspective. Nonetheless, in practice, it seems that evaluation approaches combine different elements of the positivism-constructivism spectrum (Rohracher et al., 2022). While this perspective offers important insights into TIP evaluation, I was still looking for an approach that would consider causal mechanisms and provide explanations regarding how outcomes come about.

Along the positivism-constructivism spectrum, evaluation approaches rooted in critical realism, such as realist evaluation, emerge as an alternative, as described in Section 2.2. It is from this perspective that we departed when developing an integrated framework for evaluating TIP,

introduced in Paper III. Three main reasons underlie this choice. First, realist evaluation, together with ToCs, comes as an alternative to methods-driven evaluations, such as those based on experimental and quasi-experimental methods (see Section 2.3). Such approaches allow evaluators to look at what the program actually does by exploring processes, context, and outcomes, and thus deal with two main challenges related to experimental and quasi-experimental approaches (Rolfe, 2016). One relates to the black-box problem, wherein these approaches just provide a description of the outcomes rather than an explanation of how and why programs work (Pawson and Tilley, 1997). The other relates to the complexity problem, in which such approaches assume that some elements are fixed or constant and hence unimportant for explaining outcomes. Theory-based approaches in turn argue that context matters and, as a consequence, the link between context and policies is important, as well as the processes leading to outcomes (Weiss, 1998).

Second, when it comes to addressing additionality, there is a need to establish causality. However, establishing causality is difficult for complex interventions. As mentioned above, when it comes to TIP interventions, there is a lack of a clear counterfactual; approaches to measuring it based on defining control groups are thus not suitable. An alternative is to use theory-based approaches together with the concept of additionality, in which program theory serves as a hypothesis of what the counterfactual could have been (Hind, 2010). In particular, realist evaluation, based on critical realism, offers an epistemological ground to study processes and context that relies on generative causation as an alternative to the successionist approach to causality (see Section 2.3). Additionally, while ToC is silent about its epistemological views, it "can potentially cope with long and complicated chains of causality within a program, establishing at least some evidence of which elements work, rather than being limited to the binary assessment of whether whole programs work offered by experimental approaches" (Rolfe, 2016, p. 46). Therefore, theory-based approaches can offer a way to address additionality without relying on traditional experimental methods, while still providing an alternative to account for causality.

Third, as discussed in Section 3.3, sustainability transitions frameworks resonate well with critical realism and the role of causal mechanisms in explaining "how things work or how outcomes are brought about" (Geels, 2022, p. 10). Based on this view, the author suggests using the concept of complex causality to explain socio-technical transitions. Therefore, theory-based evaluations

rooted in critical realism, such as realist evaluation, seemed to offer a way to evaluate TIP programs while accounting for the causal effects as needed in impact evaluations.

#### *4.3.2.2 Developing an integrated framework*

Based on the chosen evaluation methodology, Anna and I then developed an integrated framework to account for directionality and additionality in TIP program evaluation, which relates to RQ2 of this PhD thesis. As mentioned above, we depart from theory-based evaluation approaches, such as realist evaluation and ToC. As discussed in Section 2.3, previous work combined these approaches to address previous critiques (Blamey and Mackenzie, 2007; Rolfe, 2019). While ToC provides a very good description of the program implementation, RE offers a better conceptualization regarding causal explanation through the notion of generative causality.

We therefore combine insights from the TIP literature with theory-based approaches to evaluation to develop a step-by-step integrated framework for evaluating TIP programs. A step-by-step combination of these two approaches is further described in Paper III. Together, realist evaluation and ToC served as a basis for developing the program theory component of the integrated framework, where they aim to explicate the program's theory of change and identify potential causal mechanisms. The TIP literature works as the social science theory underlying the intervention and covers not only previous STI policy framings but also insights from the sustainability transitions literature. Accordingly, we use the TIP literature in defining the transition focus of the program (first component of the integrated framework), for example, by identifying the desired transition pathway of the intervention (Geels et al., 2016; Geels and Schot, 2007). Additionally, TIP serves as a basis for analyzing the socio-technical change processes in the second component of the integrated framework. The transformative processes covered in Paper IV are thus alternatives evaluators can use for identifying these processes (see Section 4.3.2.3). To explain causal effect relationships, we specifically rely on realist evaluation and the notion of generative causality. Following this perspective, it is possible to test and refine the program theory. This completes the system analysis component of the integrated framework and serves as a basis for evaluating the additionality brought about by the program at the system level. Furthermore, by identifying processes of change in the socio-technical system, actor networks, and institutions, the assessment of directionality is within reach. Therefore, the evaluator can check whether the desired

transition pathway defined in the program theory is being realized or not, which is one of the steps related to the third component of the integrated framework (i.e., synthesis and overall assessment). This can then lead to program revision and improvement (i.e., formative evaluation), leading to a re-start of the cycle of evaluation.

The resulting integrated framework is composed of seven steps divided into three components of TIP evaluation, that is, program theory, system analysis and synthesis, and overall assessment, and is illustrated as a cycle in Figure 1 of Paper III (see Figure 8 of this thesis). It is worth noting that the cycle might be interpreted as a linear process, but the evaluation itself should be an iterative exercise, where the different steps of the framework influence and are influenced by the others. Therefore, the cycle simply serves as an illustration of the evaluation process.

#### *4.3.2.3 Identifying transformative processes*

To analyze socio-technical system change processes, Anna and I then looked at the sustainability transitions literature to identify both the innovation- and transition-related processes the evaluators should look for, namely transformative processes. This serves as an alternative to such possibilities as the transformative outcomes proposed by Ghosh et al. (2021) to identify the processes leading to system change.

Our point of departure was a few criticisms in the sustainability transitions literature regarding the capacity of the TIS framework for explaining transitions (Markard et al., 2015), and the fact that previous work discussed the possibility to integrate the TIS and the MLP perspectives to better explain changes in socio-technical systems (Kivimaa and Kern, 2016; Markard and Truffer, 2008). Additionally, some authors pointed out that the TIS framework was elaborated based on processes of niche nurturing, and would pay little attention to "*shielding* of emerging innovation systems against mainstream selection pressures" (Smith and Raven, 2012, p. 1029).[31] Given these shortcomings, we identified a set of non-overlapping key transformative processes, while capturing both directionality and behavioral additionality at the system level.

[31] As put by Suurs and Hekkert (2009, p. 154), "if the SNM approach stops at the niche level, the TIS approach explicitly studies the growth of a niche into a regime."

As a first step, we selected a list of papers describing the processes of transformative change from the sustainability transition perspective (i.e., MLP and SNM). We then gathered a list of processes from these papers and organized a series of workshops (between Anna and me) to identify those that overlapped the TIS functions and those that should be added to the functions. Most of these processes were related to the structural elements of the innovation system, such as changes in actors, networks, and institutions (Markard et al., 2015). Others were related to processes of creative destruction and the decline of socio-technical systems (Kivimaa and Kern, 2016). This resulted in three clusters of transformative processes: socio-technical system, actor networks, and institutions. These processes represent the type of changes the evaluator should be looking for when evaluating TIP programs, reflecting the suggestion by Janssen (2019) that we should look for what has been added to the system dynamics due to policy intervention (i.e., additionality).

To address the directionality dimension, we explored the potential of the functions' framework to assess the system's dynamics by adding a "directionality filter" to each function. The reasoning behind this relates to the fact that the function can be applied at different levels of analysis (i.e., individual technologies, groups of technologies, sectors, etc.), and the directionality filter could then capture the direction of change by answering questions such as who is involved and what technologies are being targeted and experimented with. Based on a system analysis that considers directionality, then, evaluators can assess whether the transition is going in the "right" direction.

### 4.3.3 Identifying translation challenges for practice

Given that evaluation is inherently practical, looking into practice also becomes relevant from a scholarly perspective, especially because practice seems to be "lost in translation" when it comes to TIP, as discussed in Section 1.2. Hence, the third research question addressed in this PhD thesis concerns identifying the challenges related to translating an "ideal type" TIP evaluation strategy into practice. This has been addressed in two ways. First, I investigated the main challenges involved in adapting current innovation policy practice to TIP. Thereby, I examined three innovation policy programs implemented by Vinnova and how the evaluation of these programs was done in practice, that is, Vinnväxt, CDI, and SIPs, using the analytical framework for an "ideal type" TIP evaluation strategy, covered in Paper II.

As discussed in Section 4.2, this paper was prepared in collaboration with Emily and Erik as part of the GReaTr project. Therefore, the choice for looking at Vinnväxt, CDI, and SIPS was originally determined within this project. However, beyond this, these three programs represent the most recent efforts from Vinnova to develop innovation systems. Specifically, Vinnväxt focuses on the development of regional innovation systems, CDI promotes innovation and collaboration toward addressing societal challenges and aims at contributing to Agenda 2030, and the SIPs aim at developing a mix of sectoral and technological innovation systems in a set of key areas of importance to Sweden. Although these programs were not designed to be transformative from the outset, they still show some characteristics of TIP, such as their long-term focus and the involvement of multiple stakeholders. Additionally, these three programs have their own evaluation procedures and, as identified by Borrás and Laatsit (2019), Sweden is one of the few European countries to have developed systemic evaluations. Therefore, analyzing these programs' evaluation strategy through a TIP lens serves as a good way to identify the challenges policymakers can face when making evaluations more TIP-oriented.

Having selected these programs as our cases, we follow a cross-case analysis research design (Eisenhardt, 1989). A more detailed account of the data sources, collection, and analysis can be found in Paper II. In sum, we rely on a variety of sources, ranging from interviews with policymakers and practitioners involved in these program evaluations to external evaluation reports and internal documents around the call for evaluations from Vinnova. For the data analysis, we used the analytical framework categories as themes, which were manually coded by Emily and me. We also used two half-day workshops with Vinnova to complement and adapt our findings. To proceed with a cross-case comparison, we looked at similarities and differences between the programs in relation to each category of the analytical framework, as suggested by Eisenhardt (1989), and classified their alignment in relation to an "ideal type" TIP evaluation based on thresholds, detailed in Appendix A of this respective paper. Finally, we use this comparison to identify three cross-cutting challenges that, although difficult to address, can help evaluations become more TIP-oriented.

Second, I reflect on the methodological challenges evaluators may face when translating scholarly recommendations to evaluate TIP programs into practice, which is covered in Paper V. To do this, I adapt the integrated framework that Anna and I developed in Paper III and apply it to the

BioInnovation SIP as a case study (Yin, 2009). In general, a case study follows a linear but iterative process covering the planning, designing, preparing, collecting, analyzing, and sharing stages (Yin, 2009). I chose BioInnovation not only because the broader research project focuses on the transition toward a bioeconomy, but also because the program has already been scrutinized in the TIP literature (Grillitsch et al., 2019). The adapted framework relies on the same components we proposed in the original paper but integrates a few steps for analytical purposes, which results in five steps instead of seven, as specified in Table 2, which corresponds to Table 1 of Paper V. It is worth noting that the goal of the paper was not to perform the evaluation per se, but to identify the challenges that evaluators might encounter when translating TIP program evaluation frameworks to practice.

Table 2. Analytical framework for evaluating transformative innovation policy programs used in Paper V.

| Component | Steps | Specification |
|---|---|---|
| **Program theory** | 1. Define the transition focus of the program | ▪ Nature of the problem: societal challenges and/or missions<br>▪ System delineation: focal socio-technical configuration in terms of actor networks, institutions, and technologies<br>▪ Development paths: transition pathways, solution directionality, angles of directionality |
| | 2. Explicate the program's theory of change and identify potential causal mechanisms | ▪ Specify the program's ToC and hypothesize CMOc based on realist evaluation |
| **System analysis** | 3. Analyze socio-technical change processes and test and revise causal mechanisms | ▪ Analyze system-level change processes and look for evidence while testing and revising previously hypothesized CMOc |
| **Synthesis and overall assessment** | 4. Assess trajectory of change | ▪ Transition pathways, angles of directionality |
| | 5. Revisit (and revise) program theory | ▪ Revised expected outcomes |

Source: adapted from Haddad and Bergek (2023).

To perform the case study, I used a combination of secondary data and primary data. These involve (i) program description documents provided by Vinnova, (ii) program description documents from the BioInnovation website (BioInnovation, 2023a), (iii) the BioInnovation strategic agenda (BioInnovation, 2020a), (iv) projects description and final project reports obtained from BioInnovation and Vinnova's project database (BioInnovation, 2023b; Vinnova, 2021a), (v) the

BioInnovation early report (BioInnovation, 2023a), (vi) the results from previous interim evaluations of BioInnovation (Technopolis, 2020a, b), and (vii) nine semi-structured interviews conducted with BioInnovation program managers, team of experts, and Vinnova representatives between November 2022 and February 2023 (see interview guide in Appendix A).

As for the data analysis, I use a mixed methods approach. In Step 1, I use a document analysis (Bowen, 2009) of the program description documents to define the transition focus of the program. I code the data based on the transition pathways proposed by Geels and Schot (2007) and Geels et al. (2016) and introduced in Section 3.2.3. In Step 2, I use the BioInnovation's espoused effect logic as a starting point to explicate the program's ToC and interviews to clarify some points of confusion. Based on the interpretation of this espoused effect logic, I hypothesize the CMOc intended by the program based on the transformative processes from Paper IV. [32] This exercise is similar to "theory-building" process tracing (Beach and Pedersen, 2019). As for Step 3, I combine a traditional socio-technical system analysis with "theory-testing" process tracing to test and refine the CMOc. I do this by relying on the document analysis of the evidence found in the project descriptions and final reports, BioInnovation early reports, and interim evaluations, combined with insights from the interviews, while coding for the themes related to the transformative processes. In Step 4, I use the findings from Step 3 to summarize the type of change unfolding due to BioInnovation, again relying on the transition pathways to assess directionality. Step 5 in turn refers to the restarting of the evaluation process for formative purposes. Paper V provides further details on the stepwise application of the framework and methods used.

For each step of the application of the framework, I then reflect on the key challenges evaluators might face in practice in applying a TIP program evaluation framework. Indirectly, Paper V also addresses RQ2, given that it reflects critically on how policymakers can address directionality and additionality. Finally, I integrate the challenges identified in Papers II and V into translation challenges for an "ideal type" TIP evaluation.

[32] I used the following codes: knowledge development and diffusion (KDD), entrepreneurial experimentation (EE), market formation (MK), influence on the direction of search (IDS), resource mobilization (RM), legitimation (LEG), development of positive externalities (DPE), actor networks (AN), and institutions (INS).

## 4.4 Methodological limitations

This PhD thesis has a few limitations. Herein, I point out the main limitations I encountered during my PhD journey and the measures I took to counteract them. The first challenge relates to the development of the integrated framework in Paper III. While I was familiar with previous approaches to STI policy and TIP, due to the TIP course and efforts toward developing the systematic literature review, as mentioned above, one of the challenges I experienced related to integrating theory-based evaluation approaches into TIP. As mentioned in Section 4.2, while exploring a perspective that would best fit TIP characteristics to address directionality and additionality, I had to familiarize myself with evaluation methodologies besides traditional methods-oriented approaches with which I already had some familiarity. This was a learning process, which did not come without its challenges. One example is the heavy evaluation jargon, such as the concept of CMOc of realist evaluation. However, through extensive reading and discussions with Anna throughout the process of developing the integrated framework, I was able to better understand what it would mean in relation to the social science theory, that is, TIP, I was applying. Two later events contributed to my better understanding of the term. The first related to further elaborating what context, mechanisms, and outcomes would mean in connection to the TIP literature. In particular, outcomes are specifically covered in Paper IV. The second relates to my first attempt to operationalize the framework in my Licentiate thesis by examining the BioInnovation SIP.

Given the conceptual and methodological nature of the work I developed within the first years of my PhD, the next step was to implement and investigate empirically what I had proposed so far. In this process, I underwent two additional challenges. One related to the cases chosen to investigate evaluation practice based on the analytical framework my co-authors and I developed in Paper II. As described previously in this section, we used three of Vinnova's innovation programs to understand how their evaluation practices aligned with an "ideal type" TIP evaluation strategy: Vinnväxt, CDI, and SIPs. These cases, however, were not initially designed to be transformative; many readers might thus find our case selection a bit odd. However, as I argued above, these cases still represent the best examples of innovation policy programs to date incorporating some characteristics of TIP, for example, involving a broader set of stakeholders and trying to widen the programs to address societal challenges. Additionally, Sweden has been found

to be one of the few countries to develop systemic evaluations (Borrás and Edquist, 2013), which makes the country ideal for research practice in relation to the translation process toward more TIP-oriented programs. Although we do not investigate other innovation programs in other countries, we opted to investigate these cases in-depth to identify challenges for practice in Sweden, which can potentially inform other researchers and policymakers around the world. In addition, the cases selected have been scrutinized in relation to TIP before (Borrás and Schwaag Serger, 2022; Rohracher et al., 2022; Wise et al., 2022).

The second additional challenge concerns the operationalization of the integrated framework proposed in Paper III to the BioInnovation SIP case. During my efforts to apply the integrated framework, this proved to be a complex and time-consuming process. However, it also served as an opportunity to reflect on the feasibility of the framework and how evaluators can navigate this complexity while not making the evaluation a never-ending process. This issue is addressed in Paper V, in which I am critical toward the integrated framework to generate more practical insights for policymakers and evaluators, while also pointing out the main challenges they should face when evaluating potential TIP programs. Additionally, in relation to the BioInnovation SIP case, I experienced constraints in collecting both primary and secondary data. In relation to the former, I was not able to interview all the program managers involved in the program, given that some were reluctant to participate in such interviews. Regarding the latter, I had some difficulties gaining access to all final reports produced by the projects from BioInnovation, given that not all of them are made public in the two databases I used to search for such documents (BioInnovation, 2023b; Vinnova, 2023). When I contacted Vinnova, they were reluctant to provide me with the ones I could not find online due to data access and protection issues. Therefore, given that this was beyond my control, for my data analysis, I relied mostly on the reports I could find online. However, given that the goal of Paper V was to reflect on the challenges rather than focusing on the results of the evaluation per se, the data access problem did not prevent my research process from identifying the practical challenges of applying TIP program evaluation to BioInnovation.

# 5 The papers

In this section I provide a summary of each appended paper included in this PhD thesis, while discussing my main contributions to the paper and their publication status and connection with the thesis research questions. Figure 4 illustrates my PhD process chronologically and when each paper started to be developed until their submission or publication.

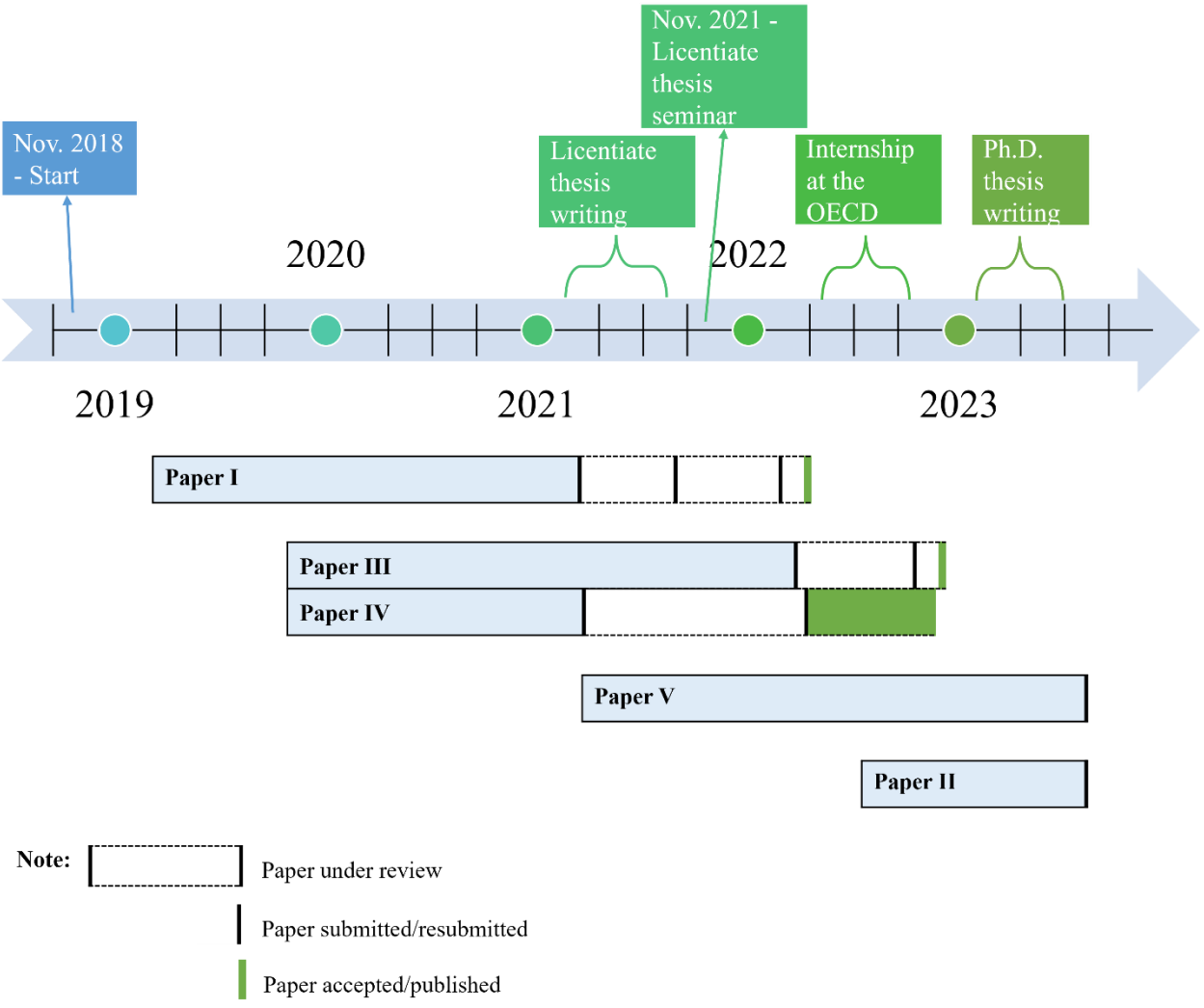


Figure 4. Overview of the research process chronology.

## 5.1 Paper I

### 5.1.1 Summary

Paper I aims to: (i) take stock of the current understanding of the specificities of transformative innovation policy and the challenges it involves throughout the policy cycle, and (ii) examine the actual contributions of the received literature to practical policymaking. To address the paper's first aim, we analyze the current state of knowledge regarding TIP in terms of its characteristics and contributions to the policymaking process through a systematic literature review of 46 academic papers. The findings of our systematic review yield insights into three main points related to the TIP state of knowledge. First, we identify five distinguishing characteristics of TIP: (i) grand challenges and inclusive growth, (ii) directionality, (iii) multi-faceted policy intervention, (iv) multiple actors and global networks, and (v) multiple levels of governance.

Second, we analyze the insights of the selected papers for the TIP policymaking process, based on a six-stage policy cycle model, including agenda-setting, policy formulation, legitimation, implementation, monitoring and evaluation, and policy learning. For each stage, we analyze how the five distinguishing characteristics of TIP could influence it and what challenges emerge according to the reviewed papers, which are referred to as "first-order" challenges. Overall, we identify an emphasis on discussing the agenda-setting stage of the policy cycle, possibly due to the focus on directionality, and a lack of attention to the legitimation stage.

Third, to address the paper's second aim, we group the identified challenges related to the different stages of the policy cycle into nine "second-order" challenges that represent challenges that cut across different stages of the policy cycle and relate to the five characteristics of TIP, as illustrated in Figure 5. Based on these second-order challenges, we identify which framework put forward in the reviewed papers explicitly bring insights to policymaking practice. We find that while some second-order challenges have been addressed in the literature more often, such as #1 and #4, others such as #5 and #9 remain largely unaddressed.

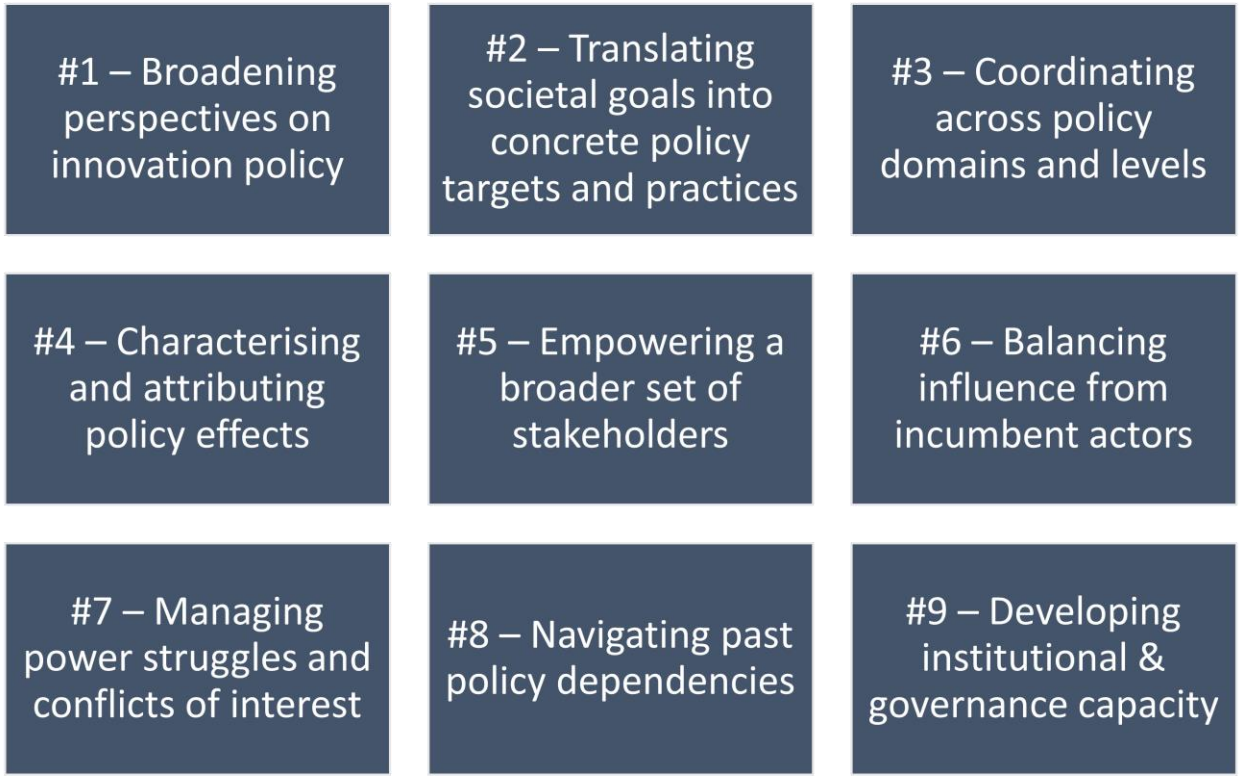


Figure 5.Second-order challenges. Source: adapted from Paper I.

### 5.1.2 My contributions to the paper and publication status

The literature study was designed and carried through a joint effort by all the authors, that is, Valentina, Anna, Hans, and I, including the choice of keywords and selection of articles to include. The search as such (including the compilation of results) was done by Valentina and me. All authors were equally involved in the analysis of the articles, in terms of both the identification of themes and challenges. All authors also participated in writing the paper. Specifically, I wrote the first draft of Section 3 (together with Hans), Section 5 (together with Valentina), and Sections 2, 4.1, and 4.4. I was also responsible for revising these sections before submission. In the revision process, I revised these same sections according to suggestions from reviewers. This also included the development of an expanded version of Section 2 and the inclusion of Appendices A and B. A first draft of the paper was submitted and presented by Valentina and me at the International Conference on Public Policy 4, held in Montreal, Canada, in 2019. The paper was then submitted to the *Environmental Innovation and Societal Transitions* (EIST) journal in March 2021. After three rounds of revision, the paper was accepted and published in March 2022 (see Figure 4).

### 5.1.3 Connection with the thesis research questions

This paper is mostly conceptual, as it develops a systematic literature review of TIP (Jaakkola, 2020) and addresses the first research question of this PhD thesis, regarding the identification of distinguishing characteristics of an "ideal type" TIP evaluation according to theory. Indirectly, Paper I also discusses some of the challenges for TIP evaluation in relation to the different TIP characteristics, but mostly from a conceptual perspective rather than practical, as intended by the third research question considered in this thesis. Some of the challenges highlighted by Paper I, such as the attribution of transformative effects is a recurring issue in policy practice, are nonetheless later highlighted in Papers II and V.

## 5.2 Paper II

### 5.2.1 Summary

Paper II aims to: i) understand what TIP means for evaluation according to the TIP literature, and ii) investigate the extent to which TIP thinking has been adopted in evaluation practice and the main challenges policymakers face in making it more transformative. To address Paper II's first aim, we develop an analytical framework based on earlier work from Edler et al. (2012), who identified four main dimensions that characterize evaluations: evaluation set-up, policy measures (merged with the first dimension), main questions, and methods.[33] By adapting this framework, we specify how an "ideal type" TIP evaluation differs from previous STI policy evaluation. This results in an analytical framework along three main dimensions, which are subdivided into categories for TIP evaluation: (i) evaluation set up, subdivided into evaluation timing, purpose, scope of the evaluation, level of analysis, and stakeholder involvement, (ii) main questions, subdivided into input/output/outcome, additionality, and directionality, and (iii) methodology, which is composed of a single category, namely causality.

Regarding the paper's aim and based on the resulting analytical framework detailing what an "ideal type" TIP evaluation strategy would look like, we analyze the evaluation strategy of three Swedish innovation programs: Vinnväxt, CDI, and SIPs, funded by Vinnova. Based on the findings from

[33] The authors identify a fifth category, related to policy and quality, but we did not include it in the analysis since these issues arise come after the evaluation itself and are not expected to differ for TIP.

the case studies, we perform a cross-case analysis to compare each program's relative alignment with the "ideal type" of TIP evaluation. We then discuss three cross-cutting challenges we identified in cross-case analysis: (i) developing (transformative) theories of change, (ii) conceptualizing systems, and (iii) addressing system-level additionality and directionality.

### 5.2.2 My contributions to the paper and publication status

This paper started out as a report to Vinnova, comparing the evaluation experiences of three of the abovementioned innovation programs. In that process, my role was to write up the case on the Swedish SIPs based on secondary materials I had gathered, and interviews organized by Emily and performed jointly by the two of us. When reworking the report into a scientific article, I took the lead in developing the analytical framework based on a less scholarly anchored version developed by Emily and Erik (Wise and Arnold, 2022). I also wrote the introduction, cross-case analysis, discussion, and conclusions, revised the methodology section, revised the SIP case (to fit the framework), and reviewed and revised Emily's draft version of the other two cases, that is, Vinnväxt and CDI (to align the three cases). Earlier versions of Paper II have been presented in a workshop organized by Susana Borrás in Copenhagen in April 2023, the GReaTr workshop organized in May 2023, and at the European Forum for Studies of Policies for Research and Innovation (Eu-SPRI) Conference 2023, held in June in Brighton, United Kingdom (UK). The paper was submitted in August 2023 to an academic journal and is currently under review (see Figure 4).

### 5.2.3 Connection with the thesis research questions

This paper includes both conceptual and empirical parts and provides contributions to RQ1 and RQ3 of this PhD thesis. Specifically, it develops an analytical framework that draws on the evaluation dimensions and categories of Edler et al. (2012) and describes how they would differ for TIP, delineating what would characterize an "ideal type" TIP evaluation (hence, contributing to RQ1). Additionally, empirically, the paper looks at three innovation programs implemented by Vinnova, as mentioned above, and provides insights into how their evaluation practices are aligned with the "ideal type" TIP evaluation, while identifying challenges for moving toward it (hence contributing to RQ3).

## 5.3 Paper III

### 5.3.1 Summary

Paper III aims to develop a framework for evaluating transformative innovation policy programs that integrates insights from the literature on TIP and theory-based evaluation to address directionality and system-level behavioral additionality. The theory-based approach composes what the literature calls the program theory underlying the evaluation, whereas the TIP literature covers the social science theory, which informs the type of change the program is trying to achieve. The framework is composed of three main components and seven steps, which are combined to address the aforementioned evaluation challenges. The three components include: (i) program theory, (ii) system analysis, and (iii) synthesis and overall assessment. Based on this integrated framework, we provide an illustrative example of the operationalization of the framework on the BioInnovation SIP in Sweden.

### 5.3.2 My contributions to the paper and publication status

This mainly conceptual paper was developed through intense discussions between the authors over a long period of time and has existed in several versions (and at one time also included what would eventually become Paper IV). In hindsight, it is a bit tricky to identify the specific contributions of each of the authors. The literature studies that the paper is (partly) based on were conducted by both authors, where I had the main responsibility for reviewing the evaluation literature and the literature on causality, whereas Anna focused on the STI policy evaluation literature and previous evaluation frameworks put forward in the sustainability transitions literature. Both authors contributed equally in writing the paper. I focused on writing the section about theory-based evaluation, the components and steps of the integrated framework, the summary, and the discussion. Anna focused on writing the introduction and the section about previous STI evaluation approaches.

Previous versions of the paper were presented in an alternative session of the Eu-SPRI Conference, held online in June 2020,[34] and at the 11th International Sustainability Transitions (IST) Conference in August 2020 (Haddad and Bergek, 2020). The paper was submitted to *Research Policy* in

[34] The original conference was cancelled because of the COVID-19 pandemic.

January 2023. During the review process, I was responsible for developing an illustrative empirical application example based on previous empirical work from my Licentiate thesis (both study and writing were done by me alone). The paper underwent an additional round of review, and, after minor adjustments, was accepted in November 2022 and published in December 2022 (see Figure 4).

### 5.3.3 Connection with the thesis research questions

Paper III brings both conceptual and methodological insights to bear to address the second research question of this PhD thesis. Specifically, conceptually, the paper integrates the literature on policy evaluation and TIP. Methodologically, the paper illustrates an iterative cycle with a number of steps evaluators can follow to address directionality and additionality in the evaluation of TIP programs. Paper III also indirectly addresses the first research question of this thesis, as it discusses how directionality and system-level additionality can be addressed in an evaluation framework and hence provides further insights into these characteristics of TIP evaluation.

## 5.4 Paper IV

### 5.4.1 Summary

Paper IV, a book chapter, aims to identify a set of non-overlapping key transformative processes that captures both directionality and behavioral additionality. To address this aim, we first broaden the concept of TIS to cover the notion of a sector-level innovation system, which encompasses different system levels, for example, sectors, technologies, or groups of technologies. This sector-level TIS captures both innovation and transition processes underlying a specific societal sector and its different directions of development (including regime recreation). [35]

Second, based on this broader notion of TIS, we delineate three sets of transformative processes: socio-technical systems, actor networks, and institutions, by considering overlaps between different frameworks advanced in the sustainability transitions literature. Additionally, we propose the addition of a "directionality filter" within each function to capture both bottom-up activities influencing the direction of socio-technical systems—as proposed by Yap and Truffer (2019)—

[35] This contrasts with the mission-oriented innovation system (MIS) proposed by Hekkert et al. (2020), which is more limited to innovation activities aimed at a given societal challenge.

and the directionality emerging from top-down government intervention. While the functions focus mainly on the emergence of new systems, less attention was paid to the structural dynamics of systems related to changes in actor networks and institutions. Based on the SNM and MLP frameworks, we list some of the processes the evaluator can look for when evaluating changes in these structural aspects. These result in four overarching transformative processes related to changes in actor networks, and another four regarding changes in institutions, both at the regime and niche levels.

Based on these three transformative processes, we argue as Janssen (2019) does that additionality can be captured by analyzing the contribution of policy intervention to the build-up of system dynamics. Therefore, these three sets of transformative processes can better capture changes across different dimensions of the socio-technical system, as well as different levels of analysis (i.e., niche and regime). Additionally, by including the directionality filter within the functions, we can potentially understand the types of reconfigurations happening in a targeted socio-technical system and whether there seems to be a transition unfolding.

### 5.4.2 My contributions to the paper and publication status

This conceptual book chapter is based on an analysis of articles describing different types of transformative processes at the niche and regime levels of sustainability transitions. Both authors participated equally in the review and analysis of the literature as well as in developing the final categorization of the identified transformative outcomes. The categorization was initially presented in a workshop organized by Katharina Schiller in 2020 (as a replacement for the Eu-SPRI Conference in 2020, which was cancelled due to the COVID-19 pandemic) and at the IST Conference in 2020, constituting part of the framework that would later form the basis of Paper III. In June 2020, the Transformative Innovation Policy Consortium (TIPC) launched a call for a Transformative Metrics workshop, which aimed at developing a new agenda for existing and new approaches to studying socio-technical change. Faced with this opportunity, Anna and I split the IST Conference paper into two parts, one focusing on the TIP evaluation framework itself and the other on the TIP outcomes. In February 2021, we were invited by the Transformative Metrics workshop organizers to submit the paper to a Transformative Metrics compilation book. Anna took the lead in restructuring the paper (from earlier conference and workshop versions), which

was later submitted in March 2021. After a long process of editing, the paper was published as a book chapter in October 2022 (see Figure 4).

### 5.4.3 Connection with the thesis research questions

This conceptual book chapter complements Paper III in accounting for directionality and additionality in TIP program evaluation and hence also contributes to addressing the second research question of this PhD thesis. Indirectly, Paper IV also addresses the first research question by providing further insights on the directionality and system-level additionality related to TIP program evaluation.

## 5.5 Paper V

### 5.5.1 Summary

Paper V addresses a gap between the theory and practice of evaluation by reflecting on the methodological challenges encountered in translating scholarly recommendations for evaluating TIP programs into practice. To this end, I adapt the framework from Paper III to perform an evaluation of the BioInnovation SIP in Sweden. Specifically, the adapted framework comprises five main steps: (i) define the transition focus of the program, (ii) explicate the program's theory of change and identify potential causal mechanisms, (iii) analyze socio-technical change processes and test and revise causal mechanisms, (iv) assess the trajectory of change, and (v) revisit (and revise) the program theory. Using multiple data sources and methods, I describe the application of the framework to the BioInnovation case while reflecting on the main challenges encountered in each step. I aggregate the eight challenges identified into three main ones: (i) addressing directionality, (ii) unpacking linear ToC, and (iii) addressing additionality.

### 5.5.2 My contributions to the paper and publication status

This paper was solely written by me, with guidance from Anna throughout the process. An earlier version of the paper was presented at the 12$^{th}$ IST Conference in October 2021, based in Karlsruhe, Germany, but held online (Haddad, 2021a). This conference paper in turn draws on the application presented in my Licentiate thesis, presented in November 2021. Another version of the paper was presented at the Final Seminar held in April 2023. Based on feedback from the discussant, the

paper has been adapted to reflect its current purpose. The paper was submitted in October 2023 to an academic journal (see Figure 4).

### 5.5.3 Connection with the thesis research questions

Paper V is both methodological and empirical. It addresses the third research question of this thesis by reflecting on the methodological challenges evaluators can encounter when applying a TIP program evaluation framework to the BioInnovation SIP case. Indirectly, the paper also contributes to RQ2, given that the application also reflects on directionality and system-level additionality.

## 5.6 Section summary

As discussed above, the papers contribute methodologically, conceptually, or empirically to this PhD thesis while addressing the research questions posed in Section 1.2. Figure 6 summarizes the connection between the papers and the research questions, in which solid arrows indicate direct relationships and dotted ones indirect relationships. In the next section, I combine the findings of the papers to answer the research questions.

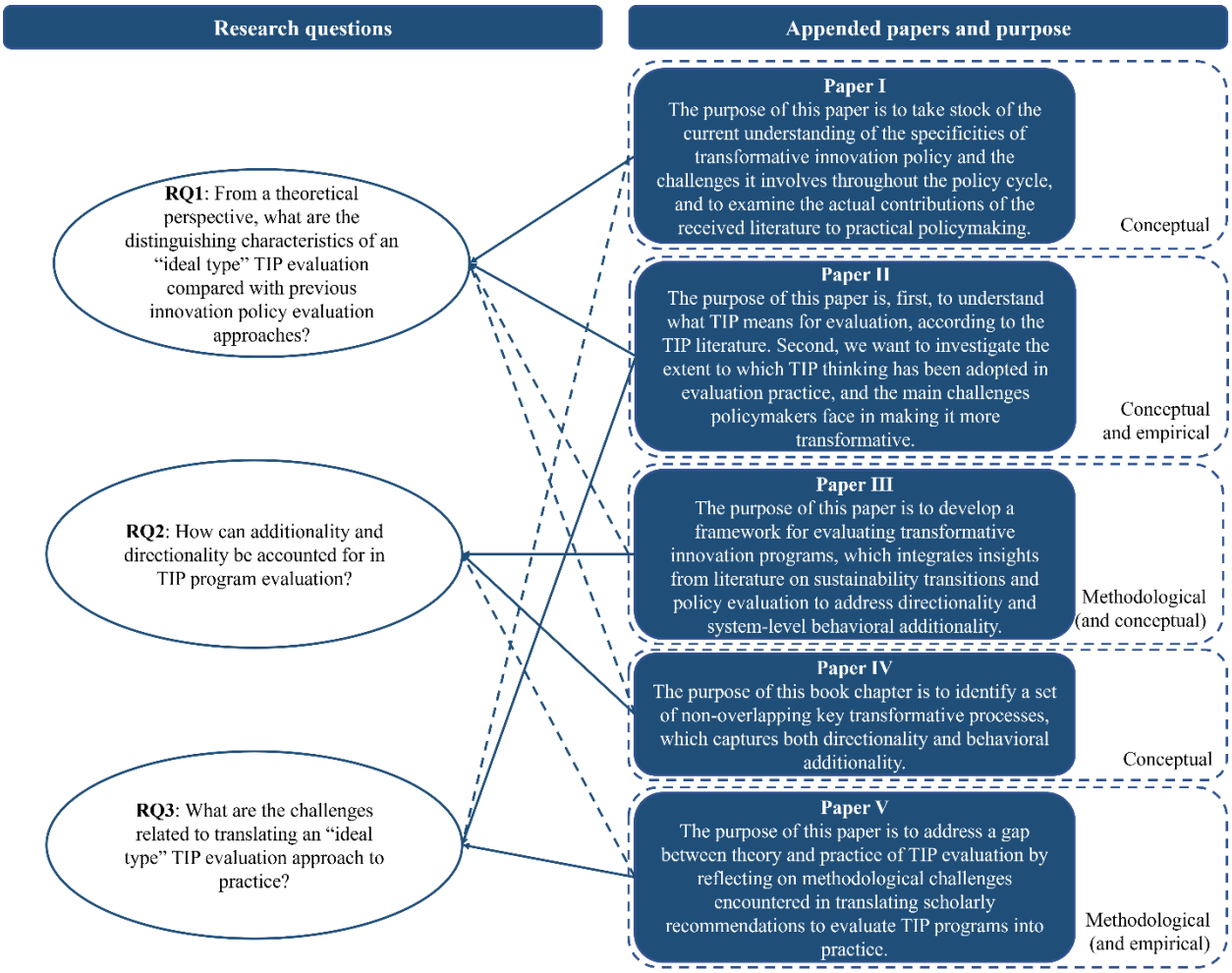


Figure 6. Connections between the research questions and the appended papers.

# 6 Findings

Below, I describe the main findings in relation to each individual paper while highlighting how they are interconnected as a step toward answering the thesis research questions.

## 6.1 TIP evaluation distinguishing characteristics

To identify the distinguishing characteristics of an "ideal type" TIP evaluation, I combine insights from Papers I and II. Paper I offers insights about TIP characteristics in general and examines how researchers refer to them in relation to the different stages of the policy cycle and, notably, to evaluation and learning. As mentioned before, in the paper we identify five distinguishing characteristics of an "ideal type" TIP: (i) grand challenges and inclusive growth, (ii) directionality, (iii) multi-faceted policy intervention, (iv) multiple actors and global networks, and (v) multi-level governance. All of these are relevant for evaluation and learning.

In relation to directionality, the paper points out that TIP presents new aims and tasks and that grand challenges and inclusive growth are complex phenomena. TIP evaluation therefore not only should measure traditional innovation outcomes, but also must account for transformative change and a broader set of impacts, as well as behavioral additionality and system-level effects. Regarding multi-faceted policy intervention, TIP involves broader policy mixes, which implies that interactions among different instruments should also be accounted for in evaluations. Concerning multiple actors and global networks, TIP suggests that evaluations should also be more inclusive and involve multiple stakeholders. Regarding multi-level governance, the literature review highlights a need for coordination between different technological fields, policy areas, and sectors, as well as for using evaluations as a governance and learning tool. This also implies the need for more formative evaluation rather than evaluation for accountability. Moreover, the uncertainties underlying TIP raise additional difficulties in attributing the effect policies and make ex-ante evaluation difficult.

The analytical framework in Paper II in turn presents how different evaluation dimensions and categories differ for TIP. Table 3 summarizes these dimensions and categories, while discussing how they differ for TIP.

Table 3. Key evaluation dimensions and categories relevant for transformative innovation policy.

| Dimensions[a] | Categories[a] | Summary |
|---|---|---|
| **Evaluation set-up** | Timing | Evaluation can be done ex-ante (before the implementation of an intervention), interim (at some point during the implementation of an intervention), or ex-post (after the implementation of an intervention) (Khandker et al., 2010). TIP evaluation tries to align different timings of evaluation into an evaluation strategy. |
| | Purpose | The purpose of evaluations can be summative (to assess the impact and establish causality), formative (to improve the intervention and support learning) (Knill and Tosun, 2012), or developmental (to adapt interventions in a learning-by-doing process) (Patton, 2010). TIP evaluation leans toward more formative and developmental evaluations. |
| | Scope of the evaluation | The scope of evaluation can involve the evaluation of single policy instruments or the entire policy mix (Rogge and Reichardt, 2016). TIP usually deals with entire policy mixes. |
| | Level of analysis | The scope of evaluation can also be at the level of projects, experiments, and programs, or involve a nested approach across multiple levels and sectors (Molas-Gallart et al., 2021). The latter is more connected to an ideal type of evaluation. |
| | Stakeholder involvement | Stakeholder involvement refers to who is involved in the evaluation, that is, R&I actors traditionally involved or a broader range of stakeholders, as suggested in the TIP literature (Haddad et al., 2022). |
| **Main questions** | Input/output/ outcome | Evaluation can focus on topics such as input or outputs of innovation, but should also cover key processes of change in socio-technical systems, that is, transformative outcomes (Bergek and Haddad, 2022; Ghosh et al., 2021). |
| | Additionality | Traditionally, additionality has focused on three types: input, output, and firm-level behavioral additionality (Clarysse et al., 2009). Some authors have also emphasized the need to cover system-level BA (Amanatidou et al., 2014), or even system-level additionalities (Haddad and Bergek, 2023). |
| | Directionality | This relates to assessing the overall contribution of innovation, where previous approaches have focused on developing innovation as efficiently and efficiently as possible, whereas the focus of TIP is on impacting the direction of innovation and change (Weber and Rohracher, 2012). |
| **Methodology** | Causality (how to address causality) | Different theories and methods are used to address causality. A traditional impact evaluation approach, based on experiments and quasi-experiments, follows the notion of regularity of succession (Hind, 2010). However, more appropriate approaches are needed when it comes to establishing causality for complex interventions, such as complex causality (Geels, 2022). |

[a] Dimensions and categories adapted from Edler et al. (2012).

By combining the insights from Papers I and II, I can list the key distinguishing characteristics of an "ideal type" TIP evaluation, which have not been considered in relation to previous framings

of innovation policy. Figure 7 illustrates the resulting characteristics. First, the **evaluation strategy covering different timing and purposes** is mostly emphasized in Paper II, which points out that evaluation should cover not only ex-post, but also ex-ante and interim evaluation and monitoring, while including different purposes of evaluation. Thus, the different evaluation timings should be complementary to support evidence-based policy, as discussed in Section 2.1. These should cover not only summative evaluation for accountability purposes, but also formative and developmental evaluation. Regarding summative and formative evaluation, Paper I particularly emphasizes that the latter should be prioritized, given the challenges for attributing policy effects to observed outcomes. Formative evaluation can help re-orient interventions and provide a more flexible approach to deal with the uncertainties underlying sustainability transitions. Developmental evaluation in turn has specifically been pointed out in Paper II, given that it allows for the continuous adaption of the intervention, which would also support reflexivity. Reflexivity in particular is also seen as a potential transformative failure due to, for example, weak leadership and lack of absorptive capacity, and serves as a rationale for policy intervention, as described by Weber and Rohracher (2012).

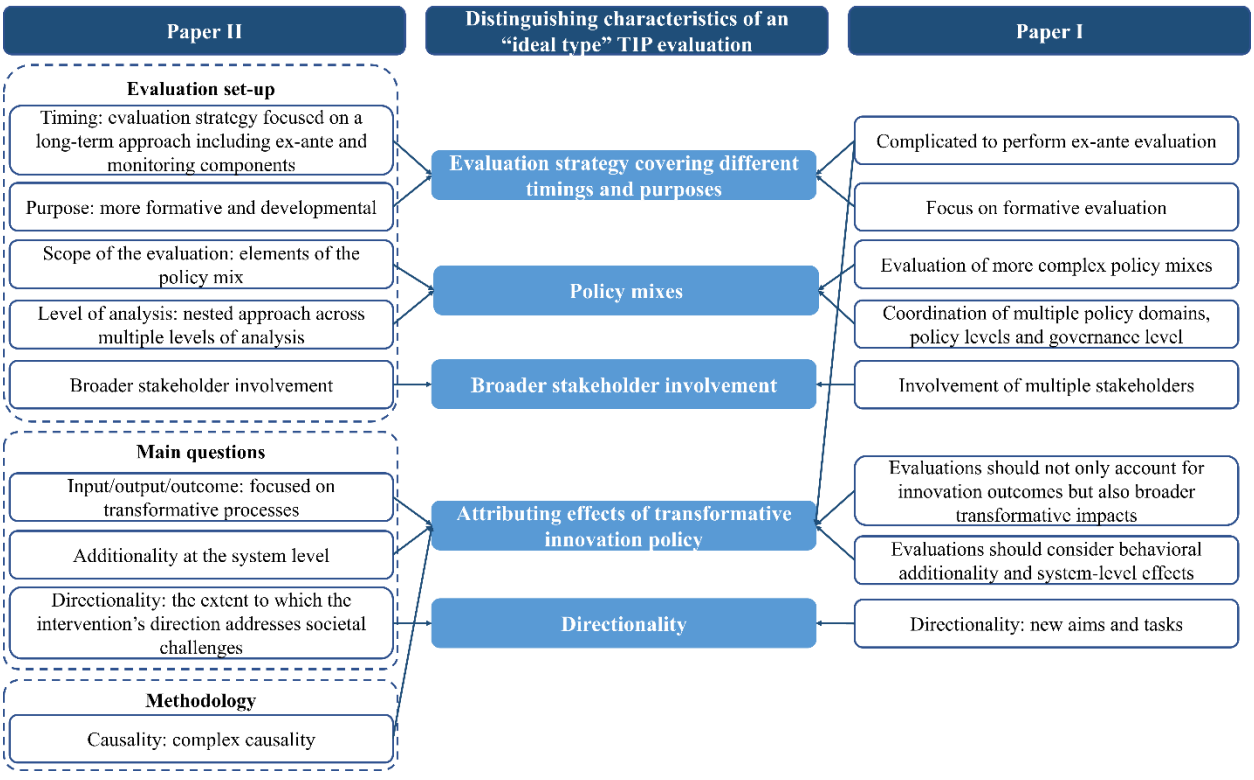


Figure 7. Distinguishing characteristics of an "ideal type" transformative innovation policy evaluation.

Second, while in the thesis I focus mostly on program-level evaluation, the TIP literature emphasizes the need to **evaluate entire policy mixes**, as pointed out in Paper I, involving both supply- and demand-side policies, but also measures to destabilize current regimes. This is because TIP is complex and involves the elements of the policy mix, that is, policy instruments and strategy, as highlighted in Paper II, however it should also account for other aspects, such as those discussed in Section 2.1. For example, how the policy process and the policy mix characteristics, such as consistency and coherence, affect the impact of policy. Additionally, other dimensions such as distinct policy fields, geography, time, and governance levels should also be accounted for in evaluations. Focusing on entire policy mixes also brings additional considerations to evaluation in terms of policy coordination, given the interaction of policies across different domains and policy levels. In Paper II, we describe the need to consider multiple levels of analysis during an evaluation in a nested approach, that is, from projects to programs. Accordingly, policy mixes would bring additional levels to be accounted for when evaluating the effects of policy. In this thesis, however, as already described in Section 2.1, I focus mainly on the evaluation of programs, which combine different elements of the policy mix rather than spanning around other aspects such as those beforementioned. This is mainly because empirically I focus on the evaluation of Swedish programs.

Third, both Paper I and Paper II stress that TIP evaluation should involve a **broader range of stakeholders**, going beyond traditional evaluation actors such as innovation agencies and program managers to embrace policymakers and other beneficiaries of the intervention, for example, service providers and civil society. This is because, as highlighted in Paper I, addressing a broader societal agenda involves actors beyond the "triple helix" in the policy process, who are able to influence and change socio-technical configurations. TIP interventions might affect and be affected by different stakeholders, and evaluations should thus be able to capture the effects of policies brought about on and by these broader range of actors.

Fourth, both Papers I and II highlight the need to **attribute the effects of transformative policy**, which is not a straightforward task because of the many feedback loops and interactions between different elements of the policy mix as well as the complex dynamics of targeted systems. This is closely related to the concept of additionality and consequently raises issues of causality and the counterfactual in the case of TIP, which are discussed in greater depth in Papers III and IV. In sum,

these papers highlight that traditional input, output, and behavioral additionality do not cover the types of changes TIP interventions should aim at. Therefore, Paper III discusses the need to consider "transformative" or system-level additionality, which comprises the transformative changes (in terms of the transformative processes covered in Paper IV) that have resulted from the evaluated policy intervention(s). In this case, traditional approaches to causality, such as the successionist view (see Section 2.2), are not appropriate. Instead, we should look for alternative views of causality, such as the complex causality proposed by Geels (2022), which goes well with the notion of generative causality from realist evaluation (see Section 3.3).

Fifth, both Papers I and II highlight the need to account for **directionality** in evaluation, that is, the extent to which the intervention addresses societal challenges, which is usually translated into terms of the intervention's aims and tasks. Papers III and IV go into this issue in greater depth regarding program evaluation, highlighting different ways to address directionality in evaluation, for example, by looking at how a program contributes to different transition pathways. Below, I elaborate further on these two last distinguishing characteristics of an "ideal type" TIP evaluation, focusing on the program level.

## 6.2 Directionality and additionality in TIP program evaluation

As discussed above, two key characteristics of TIP evaluation refer to the need to account for directionality and additionality. This is addressed in Paper III, through an integrated framework for evaluating TIP programs, and in Paper IV, through a list of transformative processes the evaluator can look for when studying the outcomes of TIP interventions. Indirectly, Paper V also covers directionality and additionality.

Regarding **directionality**, in Paper III, we suggest that directionality should be accounted for when setting up the program theory of the intervention, which refers to the first component of the integrated framework, as illustrated in Figure 8. Specifically, we argue that TIP interventions should depart from a given societal challenge or mission that should be translated into a specific development path that the program intends to stimulate. To date, there is a lack of consensus in the sustainability transition literature on how to operationalize this translation. In Paper III, we mention a few options that can be useful. One option would be to translate the targeted societal challenge or mission using the transition pathways typology proposed by Geels and Schot (2007)

and Geels et al. (2016). This would involve, for instance, reflecting on how the program envisions change regarding the different elements of the pathways, that is, multi-level interactions, actors and social groups, technologies, and socio-technical systems, and rules and institutions. I specifically tested this approach in Paper V when trying to identify challenges related to assessing directionality for the BioInnovation SIP. Accordingly, based on the analysis of BioInnovation program's description, I interpreted how the program envisioned changes in these different elements.

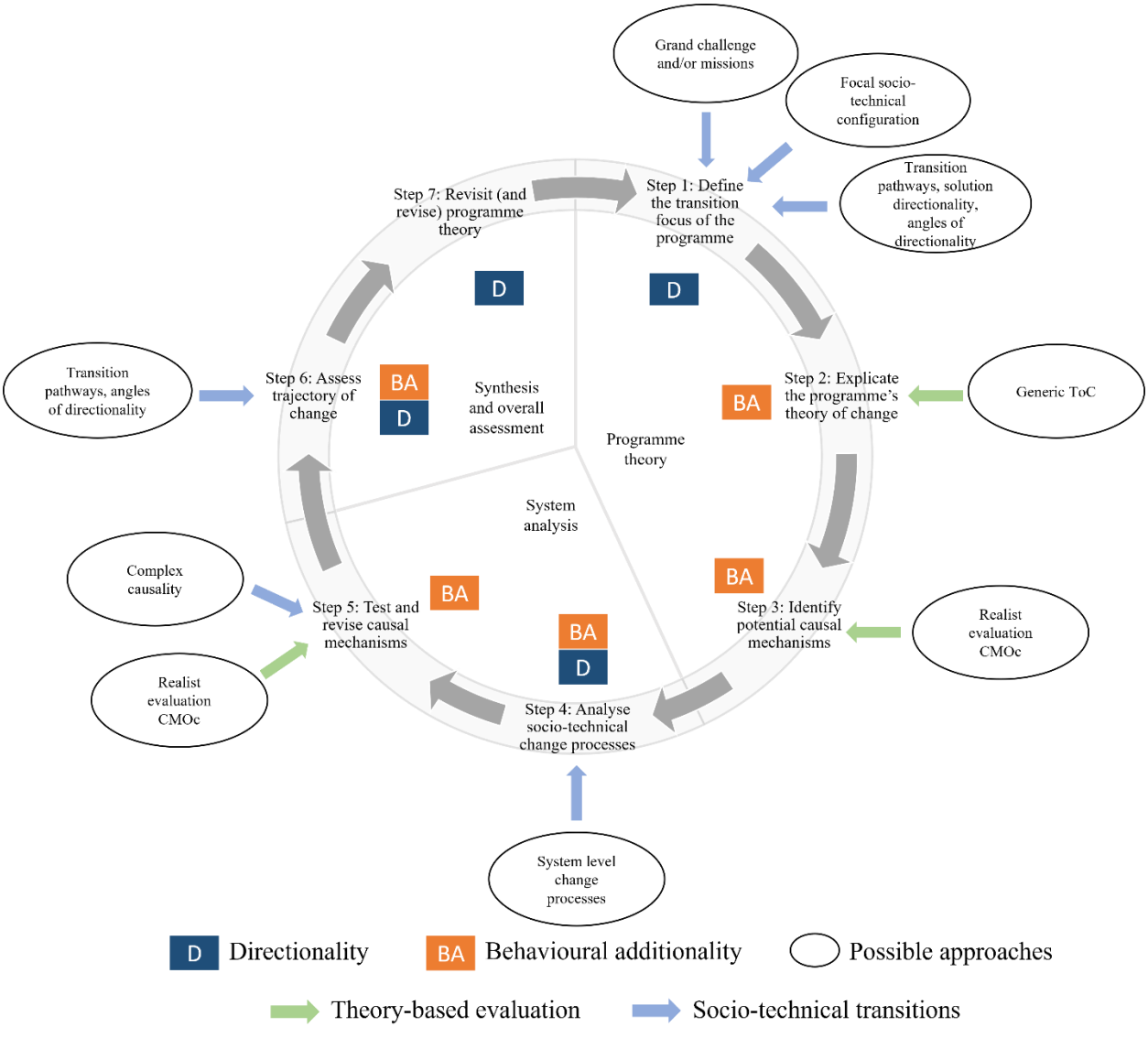


Figure 8. Integrated framework for evaluating transformative innovation policy. Source: Paper III.

Other options include angles of directionality proposed by Pel et al. (2020) and convergent problem-solution pathways suggested by Wanzenböck et al. (2020) for mission-oriented innovation policies. However, to the best of my knowledge, these approaches have not yet been tested in evaluation frameworks and hence it remains unclear how they could be operationalized for evaluation. Additionally, sustainability transition authors also seem to share different understandings of how directionality should be addressed. In particular, while Wanzenböck et al. (2020) share the idea that problem-solution pathways should converge, other authors instead point out that there could exist multiple "acceptable development paths" (Weber and Rohracher, 2012, p. 1043). Regarding this latter point, the authors point out that as building a collective vision is difficult due to the lack of consensus about the direction of change, it might be better to have a vision and let different development paths emerge. Other authors seem to argue about directionality as Weber and Rohracher (2012) do, such as Schot and Steinmueller (2018) and Köhler et al. (2019), who also point out that actors might disagree about the direction of transitions given the multitude of opinions. I also lean toward these latter understanding.

Having set the intended development path of the program, the evaluator should then assess the direction of change that emerges due to the intervention. We thus suggest in Paper III that evaluators should make a socio-technical system analysis, by looking at both innovation and transition change processes. The sustainability transition literature has also put forward a few approaches that could be used for this purpose. Specifically, in Paper IV, Anna and I propose widening the TIS functions (see Section 3.2.4) to cover key socio-technical change processes (or transformative processes), while also making the directionality more explicit. Specifically, we list three clusters of transformative processes related to changes in actor networks, institutions, and the socio-technical system (the resulting list of processes can be found in Paper IV). We argue that by identifying the changes in actor networks and institutions, we can assess the degree of change in these structural components of the system, both at the niche and regime levels. Additionally, we point out that to account for directionality, previously connected exclusively with the "guidance of the direction of search" function, we need to identify how different emerging or established technologies (or both) are being supported. We thus add a "directionality filter" to each function to assess the dynamics of different technologies more explicitly. As a result, the analysis of changes in these configurations can be used to "diagnose" whether the transition is moving toward

the intended pathway. This analysis can, for instance, be used to assess the intended transition pathway that is unfolding due to the program.

Alternative frameworks include the transformative outcomes by Ghosh et al. (2021) and, more recently, the set of innovation system functions adduced by Elzinga et al. (2023) to assess missions.[36] The former framework, as mentioned in Section 1.1.2, proposes a set of 12 transformative outcomes to guide the evaluation of TIP. These are based on a review of sustainability transition frameworks, such as the MLP. Elzinga et al. (2023) in turn adapt the TIS functions to propose missions' specific system functions, which are divided between programming and performance functions. The former functions refer to the processes concerning the mission's programming structure or mission arena (Wesseling and Meijerhof, 2021), which relate to setting directions, that is, problem and solution directions, and coordinating the governance of the mission. The latter functions in turn include six processes related to the innovation (i.e., knowledge development, knowledge diffusion, entrepreneurial experimentation, market creation, resource mobilization, and creating legitimacy) and destabilization side (i.e., unlearning, knowledge network break-down, restriction of experimentation, resource withdrawal, and challenging status quo). Despite this diversity of frameworks, they all aim at assessing processes leading up to a transition. It is thus up to evaluators to define which framework to use, preferably based on their experience and familiarity. When analyzing such processes in Paper V, I used the transformative processes that Anna and I proposed in Paper IV.

Based on the socio-technical system analysis, the evaluator can make an overall assessment of the directionality of the program, in comparison to the targeted development path defined in the program theory. This implies that the same framework used previously to define the intended development path should be used at this stage for the assessment. In Paper V, I thus use the findings of the socio-technical analysis to identify the changes in multi-level interactions, actors and social groups, technologies and socio-technical systems, and rules and institutions that can be observed. I also show that, due to the long-term of transitions, it can be difficult to identify the type of pathway that is unfolding, as will be discussed in the next section below.

[36] Elzinga et al.'s (2023) framework has not been included in Papers III or IV, as it was only published later.

Finally, the overall assessment should be able to guide a formative evaluation by providing the basis to revisit the program theory and its intended directionality. The evaluation findings can be used for learning and adaptation of the program.

Concerning **additionality**, I argue that for TIP, the traditional input, output, and behavioral additionality, as discussed in Section 1.1.2, were not able to cover the type of system-level effects that evaluators should be looking for when evaluating interventions targeting transformative change. In Paper III, we specifically emphasize that BA was not very well adapted to either innovation systems or transitions. This also follows previous criticisms from the literature regarding how BA was operationalized in STI evaluation practice. Notably, Gök and Edler (2012) showed that STI policy evaluations could not capture BA in full due in part to a lack of clarity regarding what it really is. This has also been noted by Amanatidou et al. (2014), who highlight that, even though BA was associated with the evolutionary perspective of innovation (see Section 1.1.2), in practice BA was studied based on comparative statistics at the firm level. In Paper III, we then point out that evaluators should consider key processes influencing changes in socio-technical systems, that is, those highlighted in Paper IV. A similar argument had earlier been made by Janssen (2019), who proposed that the real "bangs" evaluators should be looking for are related to how interventions contribute to building up TIS. Based on these insights, we saw the need for a new conceptualization of additionality that captures system-level effects, that is, "transformative" additionality or "system-level" additionality (although throughout Paper III, we still refer to it as behavioral additionality).

A broader conceptualization of additionality involves challenging the traditional counterfactual approach based on "sole attribution," underpinned by the successionist view of causality and traditional methods such as RCTs and quasi-experiments. Thereby, in Paper III, we propose the use of the realist approach to evaluation, which follows the notion of generative causality (see Section 2.3), combined with the ToC approach. We use elements from these theory-based approaches in different steps of the integrated framework (see Figure 8). We argue that these approaches can be used as an alternative to building the counterfactual, following previous insights from, for example, Hind (2010). In this way, rather than proving attribution definitively, the evaluator will look at what worked for whom and in what circumstances, that is, how mechanisms work in a given context to generate the outcome (or CMOc, in the realist jargon). In this way,

evaluators can look at what the program initially envisioned in terms of outcomes and analyze how the program is contributing to them. In Paper III, we thus reflect on what CMOc would mean in relation to TIP and highlight different methods that can be used for analysis. Using multi-methods also allows for complementary approaches to causality, that is, complex causality. This has been highlighted by Geels (2022) as key to building causal explanation in sustainability transitions (see Section 3.3).

In Paper V, I use process tracing to operationalize the analysis of CMOc for the BioInnovation SIP, departing from the originally espoused effect logic pre-defined by the program. This, however, has not proven easy and, among other things, generates very complex non-linear connections between mechanisms and outcomes. Additionally, many of the outcomes identified were unintended in relation to the espoused effect logic and what had been expected from the program from the outset. I further elaborate on the challenges underlying the translation of TIP evaluation to practice in the next section. Nonetheless, tracing processes and identifying how they generated their outcomes provide many insights that can inform how the program contributes to the dynamics of socio-technical systems and why. Additionally, it shows the program's weaknesses and why they occur, which can be very important for formative evaluation and can lead to program adaptation and reflexivity.

## 6.3 Translating challenges of TIP evaluation to practice

To identify the translation challenges of an "ideal type" TIP evaluation to practice, I combine the findings of Papers II and V, as illustrated in Figure 9. In Paper II, we identify three overarching challenges resulting from a cross-case analysis of the evaluation of three Swedish innovation programs, that is, Vinnväxt, CDI, and SIPs, which are based on key TIP evaluation dimensions and categories, as listed in Table 3 in Section 6.1. These challenges include developing (transformative) theories of change, conceptualizing systems, and addressing directionality and system-level additionality.

Paper V identifies similar challenges related to the translation of an "ideal type" TIP evaluation to practice, which are identified in relation to the evaluation of the BioInnovation SIP, which is used as an illustrative case. In sum, the paper lists three main challenges resulting from the aggregation of eight challenges identified through the evaluation of the program. The first challenge refers to

addressing directionality, the second to unpacking linear ToC, and the third to addressing additionality.

The challenges identified in Papers II and V then bring us to four challenges for translating an "ideal type" TIP evaluation into practice, as illustrated in Figure 9. The first challenge is **accounting for directionality** in TIP evaluation. Both Papers II and V point out that this is not trivial and depends on the type of socio-technical system change the intervention envisions, which is closely related to its conceptualization of systems. In Paper II, when analyzing the three Vinnova programs, we find that they differ in their conceptualization of systems, ranging from a narrower view of innovation in CDI to regional innovation systems in Vinnväxt and a mix between sectoral and technological innovation systems in the SIPs. While using a different understanding of systems is not a problem *per se*, adapting programs to be more transformative implies a broader view of systems. For TIP specifically, this means targeting transitions of entire socio-technical systems, as defined in the sustainability transitions literature and highlighted in Paper I. In Section 6.2, I also noted that TIP authors seem to have different understandings of directionality, which can also impact how directionality is addressed in evaluations. While in Paper V I use the transition pathways typology to operationalize the analysis of directionality, other approaches to directionality need to be tested empirically to better inform evaluation practice.

The second challenge relates to **developing transformative ToC**. In Paper I, we identified that translating societal goals into concrete policy targets and practices is one of the challenges policymakers face regarding TIP. Papers II and V both confirm that this process of translation is difficult. Accordingly, both papers show that current innovation programs rely mostly on linear-effect logics and neglect important aspects of a ToC, such as cause-and-effect relationships, unexpected outcomes, and interactions. This affects the evaluation, as it provides little guidance to evaluators for identifying expected causal mechanisms and understanding how the program is supposed to work and why. Nonetheless, as argued in Paper III, TIP evaluation should ideally depart from a program theory that specifies the transition focus and the theory of change and provides an overview of intended mechanisms and expected outcomes of the program, which should be supported by an underlying policy theory. In Paper II, we show that these are not considered in the three Vinnova programs. which mainly rely on effect logics that do not account

for cause-and-effect relationships. Therefore, developing more transformative ToC is not straightforward and remains a challenge for TIP evaluation.

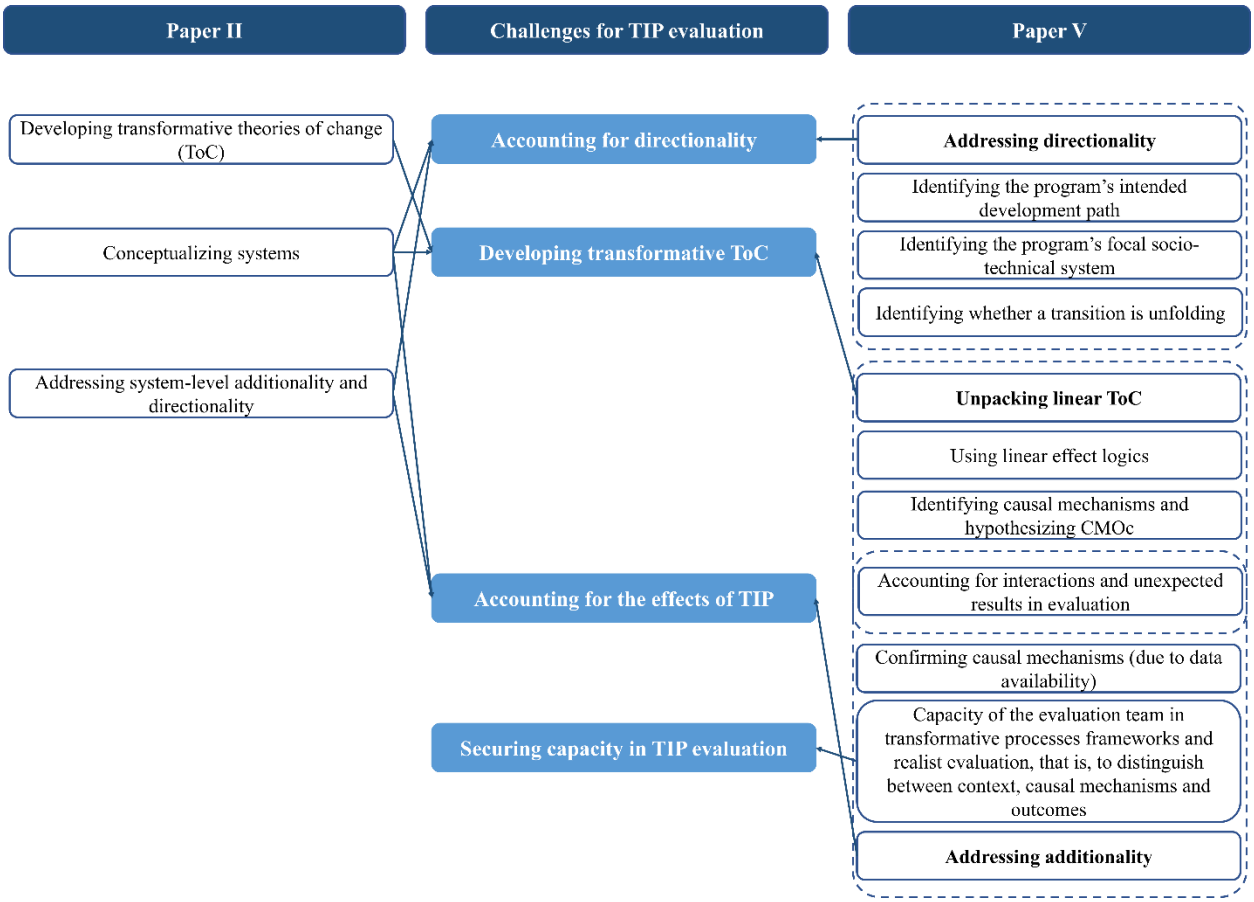


Figure 9. Challenges in translating an "ideal type" transformative innovation policy evaluation to practice.

The third challenge is that of **accounting for the effects of TIP**, which is closely related to the concept of additionality, especially in evidence-based policy. This aspect is recurrent in TIP evaluation and was already pointed out in Paper I, which showed that only a few frameworks in the TIP literature had so far addressed this. While Papers III and IV suggest a way to address this for program evaluation, Papers II and V show that it is challenging to apply this in practice. In Paper II, we showed that none of the Vinnova innovation program evaluations accounted for causal effects in practice, as they only included an overall analysis based on more quantitative indicators or processes that resemble the TIS functions. Moreover, accounting for the effects of TIP is also closely related to how the program conceptualizes systems, which underlies the whole process of evaluation. For example, Vinnväxt's evaluation sought to capture ripple effects related to the development of regional innovation systems over time, thus reflecting the type of systems change

targeted from the start. In contrast, CDI follows the definition provided by Miedzinski et al. (2019, p. 6) of a system innovation as "a portfolio of interdependent and mutually reinforcing innovations which together have a potential to transform systems delivering services to societies, such as health, food, or mobility." This narrower view of innovation also reflected the way CDI evaluation looked at indicators and its dimensions to system change. As mentioned above, however, these evaluations do not consider causality formally.

In Paper V, the BioInnovation SIP targeted system level changes, but departs from a mix between technological and sectoral innovation systems. This view also reflected the types of processes that I was able to capture when developing the program's evaluation, which are most related to the development of innovation systems. An additional point identified in relation to assessing the effects of policies in Paper V regards the data needed to support evidence tests, which were not always available for the BioInnovation case. As a result, capturing the effect of policies might also demand additional data that are not available and would trigger additional data collection efforts, which might not be feasible given time and budget constraints.

Finally, the fourth challenge relates to **securing capacity in TIP evaluation**. Paper I had already raised the issue of developing institutional and governance capacity throughout the TIP policymaking process, which we found to be widely unaddressed in the literature. In Paper V, then, I was able to better specify the specific challenges evaluators might encounter when securing this capacity in TIP evaluation. I thus argue that to perform an "ideal type" TIP evaluation, the evaluation team should be familiar with both the policy theory underlying the evaluation and the evaluation methodology used, which would entail familiarity with both the transformative processes and theory-based evaluation approaches, such as ToC and realist evaluation. This is because during the process of evaluation, the team should be able to both identify unexpected results of the program in terms of transformative processes and distinguish the context, mechanisms, and outcomes. This latter aspect has already been discussed in the realist evaluation literature as not being trivial. I also found this to be a challenge during the BioInnovation SIP evaluation. For example, when developing the socio-technical system analysis and looking for CMOc, sometimes it was not clear whether a process was part of the mechanisms, an outcome, or a contextual factor. Therefore, securing capacity in TIP evaluation represents a challenge when

seeking an "ideal type" TIP evaluation while trying to fulfil the criteria pointed out in Paper II and summarized in Section 6.1.

Taken together, the four challenges for translation are interconnected. For example, a lack of a clear and more transformative-oriented ToC will influence the type of changes and outcomes, and hence the effects of policy, that evaluators will be looking for. Developing more transformative ToCs should be informed by transformative processes, which form the policy theory underlying TIP programs. This in turn is closely related to how the program is set up and how it conceptualizes systems. In this thesis, I pointed out that TIP programs should target changes in entire socio-technical systems as conceptualized in the sustainability transitions framework, which entails a broader understanding of systems than one focusing on spatially or cognitively delineated innovation systems or on isolated innovations. This is also closely related to directionality, given that in TIP, this would involve setting a direction of change toward new and more sustainable socio-technical systems. Therefore, this also affects how directionality is accounted for in evaluations. Moreover, all these aspects require capacity of the evaluation team, not only regarding familiarity with the underlying policy theory informing the program, but also how to perform evaluations that rely on theory-based approaches, such as ToC and realist evaluation. In Section 7.3, I discuss and provide some suggestions for how to address these challenges in TIP evaluation.

# 7 Discussion

Having outlined the main findings of the thesis in relation to the research questions, I noted certain points that merit further discussion before concluding this thesis. These have three main aspects: (i) addressing the distinguishing characteristics of an "ideal type" TIP evaluation, (ii) reflections on the integrated framework for TIP program evaluation, and (iii) remaining issues of TIP evaluation.

## 7.1 Addressing the distinguishing characteristics of an "ideal type" TIP evaluation

In Section 6.1, I identified a set of five distinguishing characteristics of an "ideal type" TIP evaluation: (i) an evaluation strategy covering different timings and purposes, (ii) policy mixes, (iii) broader stakeholder involvement, (iv) attribution of effects of transformative policy, and (v) directionality. Throughout the thesis, I addressed some of these characteristics, especially (i), (iv), and (v). In contrast, others such as (ii) and (iii) need further attention. In this section, I discuss the former, whereas the latter will be discussed in Section 7.3.

First, regarding the *evaluation strategy covering different timings and purposes*, I consider it mostly in Paper II when analyzing the evaluation of the three Vinnova programs. The integrated framework from Paper III is, however, silent in relation to the evaluation timing in which it should be used, but we emphasize that it can be used for formative purposes. This latter aspect is particularly highlighted in Step 7 of the framework (or Step 5 in Paper V), which focuses on revisiting and revising the original program theory to update the program based on what is learned from the evaluation. The TIP literature seems to converge toward an agreement that TIP evaluations should be formative, as many scholars have already pointed out (Ghosh et al., 2021; Molas-Gallart et al., 2021; Rohracher et al., 2022). An alternative to formative evaluation, that is, developmental evaluation, has also been put forward in Paper II as a possibility for TIP. Developmental evaluation emphasizes the involvement of evaluators in successively improving the intervention, focusing on "adaptive development" (Patton, 2010). The author argues that this approach is especially suitable for complex interventions in which the knowledge base is underdeveloped. However, this remains to be tested in TIP evaluations.

Despite its silence on timing of the evaluation, the framework of Paper III can potentially inform other evaluation timings. For example, the development of the program theory, which is the first component of the framework, can be used for ex-ante evaluation. Thus, the evaluator can perform an ex-ante evaluation of the program theory by checking the "proposed design for adequacy of cause-and-effect reasoning as shown in program theory and [its] likely appropriateness, effectiveness, and efficiency" (Funnell and Rogers, 2011, p. 419). According to these authors, this can help identify points that need to be modified before the program is implemented. Concerning monitoring, the development of the program theory can ex-ante indicate some of the indicators one can monitor during implementation. This can also serve the purpose of establishing baseline data (Funnell and Rogers, 2011). After the program's implementation, these indicators can be monitored to inform revisions of the programs and support formative evaluation. This can also serve as a venue for future research related to the development of new monitoring approaches for TIP. In terms of ex-post evaluation, the integrated framework can still provide inputs to post-assessments of the program. For example, it can serve as a basis for summarizing the main goals achieved and whether the program was appropriate.

Evaluation can also be developed for different timings with different purposes. In Paper III, we point out that transformative change usually goes beyond specific socio-technical systems, which raises additional challenges when scoping the evaluation and defining its boundaries. We thus recommend, that evaluations can either consider a broader approach to account for multi-system interactions or a narrower one in which the evaluator focuses on specific parts of the system. Therefore, using complementary approaches to evaluation might be ideal for TIP, considering a different timing and purposes to cover different systemic levels. This is closely related to the idea of developing an evaluation strategy covering activities of monitoring, evaluation, and learning over time, as emphasized in Paper II. An evaluation strategy can thus cover different levels of abstraction, sometimes zooming out and grasping multi-system interactions across a portfolio of programs or zooming in on specific parts of the program. A similar argument has also been put forward by Janssen et al. (2022), who tentatively identified six categories of evaluation of innovation and transition programs within a summative-formative spectrum. According to the authors, these categories can complement each other in gathering different pieces of evidence about different parts of the system. The development of an evaluation strategy should also be

considered in the design and implementation of an intervention, as suggested by Molas-Gallart et al. (2021).

Third, concerning the *attribution of effects of transformative policy*, the thesis proposes an alternative way to address this for impact evaluation, which is highlighted in Papers III, IV, and V. In contrast with previous approaches of attributing the effect of policy through the counterfactual using traditional methods such as RCT and quasi-experiments, which are based on the notion of regularity of succession, I propose using alternative approaches to address causality. As argued in Section 6.2, this would encompass using realist evaluation and identifying what worked, how, and in what circumstances, that is, CMOc. Moreover, program theory would work as an alternative to the counterfactual, in which evaluators look for the contributions accomplished by the program in relation to what it envisioned from the outset. This could then allow evaluators to gain a closer understanding of the additionality brought about by the program at the system level. Despite an attempt to operationalize this in Paper V for BioInnovation using process tracing, we need more empirical applications of such an approach. Ideally, this would also involve considering complex causality and applying multi-methods to collect different pieces of evidence related to the effects of the program, as suggested by Geels (2022). This, of course, should be weighted in relation to the evaluation time and budget constraints. This is why it might be worth considering developing a more in-depth causality analysis for specific parts of the program while using other less resource consuming approaches when considering multi-level interactions, as discussed above.

Fourth, regarding *directionality*, I propose ways to address this characteristic in evaluation in Papers III, IV, and V. My argument is that from the beginning, policymakers should scope the trajectory of change the program intends to contribute to, as highlighted in Step 1 of the integrated framework from Paper III. During the evaluation, the evaluator can spot how the program is contributing to the emergence of this intended trajectory of change by analyzing transformative processes, such as those suggested in Paper IV. While there are many different approaches to addressing directionality in the TIP literature, such as transition pathways (Geels et al., 2016; Geels and Schot, 2007), angles of directionality (Pel et al., 2020), and problem-solution directionality (Wanzenböck et al., 2020), there is currently no agreement in the TIP literature on how to address directionality. Therefore, it is up to evaluators to select the approach that they deem more suitable.

In Paper V, I use the transition pathways to discuss the trajectory of change that seems to be unfolding with respect to BioInnovation. Despite the challenges of assessing directionality due to the long-term character of transitions, this approach was able to provide a general understanding of the trajectory of change to which the program seemed to be contributing. This is interesting for learning aspects regarding how a program can support a transition. Nonetheless, in Paper V, I conduct the analysis mostly based on a socio-technical system analysis. Other approaches to assessing transition pathways could also be used to complement the directionality assessment, such as initiative-based learning and quantitative system modelling, as suggested by Turnheim et al. (2015).

Other approaches could have been used in Paper V, instead of the transition pathways typology. Although they have not been tested yet, we can still conjecture how they could be operationalized in the integrated framework. For example, taking a problem-solution directionality perspective to evaluate the BioInnovation SIP could have involved analyzing what types of problems emerge when trying to transform Sweden into a bioeconomy by 2050, as well as the main solutions that could be available to address this. Following Wanzenböck et al. (2020), this approach could then serve as a diagnostic of whether problems and solutions converge or diverge and in which problem-solution space the program can be situated. This could then inform the types of activities that the program could target to follow a more convergent pathway. During the evaluation, evaluators could thus examine how the pathway is unfolding and whether the program is following the desired pathway. In the case of divergence, a formative evaluation could provide the opportunity to adapt the program and direct it toward convergence.

Alternatively, if the evaluator opts to use the angles of directionality approach, as argued by Pel et al. (2020), assessing directionality can involve a combination of different evidence, taking into consideration three angles of directionality, that is, socio-technical multiplicity, appraisal diversity, and process dynamics. For the BioInnovation case, using these angles to assess directionality could involve, first, an assessment of how actor coalitions are influencing the direction of the program over time and what types of changes are being pursued. Second, it could include an assessment of different normative views regarding what is considered relevant and what types of transition should be prioritized. Regarding the last angle, this could involve an assessment of what type of innovation trajectory can be experienced or anticipated. The assessment of these three angles,

however, might demand more planning and time from evaluators, as it considers different analytical dimensions, which might require additional efforts into data collection and the use of multi-methods at different points in time to capture the different angles. Therefore, while such an approach might yield valuable insights regarding directionality, it might be unfeasible in practice considering the resources allocated to evaluation.

Given the variety of approaches that can be used to account for directionality in TIP evaluation, it is hard to suggest which would be most suitable. Evaluators might also have different opinions regarding the best approach. Therefore, it is up to the evaluation team to define which alternative better fits the type of evaluation that they will perform and what would be feasible given the resources allocated for the evaluation itself. In terms of research, it would be interesting to test these alternatives when assessing directionality empirically. For instance, researchers could develop a comparative study of different approaches to assess directionality to identify their advantages and disadvantages.

## 7.2 Reflections on the integrated framework for evaluating TIP programs

The development of the framework (Papers III and IV) and its application (Paper V) lead me to reflect on the feasibility of the framework for performing TIP evaluation program in practice. Thereby, three main points can be highlighted in relation to: (i) additionality, (ii) directionality, and (iii) other empirical applications of the framework.

First, concerning additionality, in Paper III, we introduced the concept of system-level additionality. This concept is proposed to cover a broader view of additionality beyond the input, output, and behavioral types. While developing the work on the integrated framework and analyzing the types of additionalities that had been advanced in relation to the previous framings of innovation policy, we realized that the type of effects expected from TIP programs should be related to innovation and transition processes. These include, for instance, the transformative processes proposed in Paper IV or other alternative approaches that look at innovation and transition processes (Ghosh et al., 2021; Kern, 2012; Kivimaa and Kern, 2016). Previously, Janssen (2019) had pointed out that the type of additionality policymakers should be looking for was those related to supporting the fulfillment of the TIS functions. Accordingly, the traditional

concepts of additionality seemed too underdeveloped to account for the supposed changes TIP programs should be contributing to.

In Paper V, I then attempt to show how to use this concept in practice. Such an exercise, however, proved complex. On the one hand, it was possible to trace some of the causal mechanisms emerging due to the BioInnovation SIP, which can shed light on the additionality of the program. I was also able to identify some of the unexpected outcomes that emerged due to the intervention, which had not been accounted for directly in the program's espoused effect logic. On the other hand, I could not provide further evidence to confirm the causal mechanisms, mainly due to data constraints. Consequently, this leads me to two additional reflections.

One relates to the fact that evaluators might need to prioritize a "scientific and pragmatic" approach to causal inference, as put by Funnell and Rogers (2011, p. 469). The authors point out that "a scientific approach to causal analysis does not imply the use of any particular research design or type of data. Rather, it refers to an approach that is systematic, draws on a range of evidence, and critically reviews and synthesizes this evidence" (Funnell and Rogers, 2011, p. 471). This supports the theory of change approaches that promote the use of mixed methods in evaluation. By pragmatic, the authors argue that evaluators should take a practical approach to causal inference that is adapted to the resources available to perform the evaluation. This also involves calculating the risks of running into Type I (i.e., concluding wrongly that the intervention had an effect) and Type II errors (i.e., concluding wrongly that the intervention had no effect). The evaluator might need to balance the development of an evaluation that can reasonably look at the program theory and do minimal checking for alternative explanations against one that demands a careful construct of the counterfactual and the systematic investigation of alternative explanations. In all cases, however, the authors emphasize that causal inference needs to be considered from the beginning.

My second reflection concerns the consideration of causality from the start, that is, when designing the intervention itself. In the TIP literature, many authors have pointed out the need to align the design of interventions and implementation with monitoring and evaluation (Molas-Gallart et al., 2021; Rohracher et al., 2022). The lack of consideration of causality within the ToC, in particular, has been pointed out as one important challenge in TIP evaluation practice, as identified in Papers II and V and further considered in Section 6.3. However, the literature still does not provide much

guidance on how to effect this in practice, which also indicates that more research is needed. One exception is the Motion Handbook on how to develop transformative ToC (Boni et al., 2021) using the transformative outcomes from Ghosh et al. (2021), developed by TIPC. Additionally, as highlighted in Paper V, other frameworks such as the motors of innovation proposed by Suurs and Hekkert (2012) could potentially be adapted to TIP to support the development of transformative ToC.

Second, regarding the assessment of directionality, identifying the overarching trajectory of change and whether a transition is unfolding might not be feasible during a program's lifespan. In the SIP case, for example, the duration of the programs can be up to 12 years. In contrast, as the sustainability transition literature highlights, transitions are long-term and might take decades to unfold (Geels, 2004). During the application of the framework in Paper V, I pointed out that although having an overarching idea of how BioInnovation is contributing to the transition pathways is possible, it is difficult to know whether a transition is unfolding.

On the one hand, the assessment can still suggest some policy recommendations. In the BioInnovation case, this would be targeted at the institutional focus of the program, which seems to be following a reproduction pathway. Therefore, from a TIP perspective, further efforts toward institutional change and destabilization could be interesting. On the other hand, regarding the change in technologies and socio-technical system of the transition pathways, the results are more uncertain. This could be due to time, given that due to the long-term characterize of transitions, it could still be too early to assess how the program is supporting technological change and to spot the trajectory that is emerging as a result. Alternatively, this could be related to the methods I used in Paper V for assessing the pathways, which is based on a socio-technical system analysis. Other methods might be needed in addition to the system analysis. For example Turnheim et al. (2015) suggest using additional methods for evaluating transition pathways, consisting of an integrated approach involving quantitative modelling, socio-technical system analysis, and initiative-based learning, with the goal of covering different aspects of sustainability transitions. This, however, would not only bring additional complexity to the evaluation, but would also demand more resources and time, which might preclude such an evaluation. Yet other approaches for assessing directionality might need to be tested, for example, angles of directionality (Pel et al., 2020) or

problem-solution directionality (Wanzenböck et al., 2020), or, indeed, new approaches might need to be developed.

Finally, in relation to the application of the framework in Paper V, one empirical case is still quite limited for testing the full capacity of the integrated framework for evaluating TIP, as already mentioned in Section 7.1. Therefore, future empirical applications might be considered to test it further, including, for example, the use of the integrated framework for evaluating other SIPs or even other interventions outside Sweden to compare different empirical contexts and cases. This could also be related to testing emerging approaches to assess directionality, attempts to operationalize the concept of complex causality, or even simplifying the framework even further.

## 7.3 Remaining issues of TIP evaluation

While in Section 7.1 I discussed how this thesis has addressed some of the distinguishing characteristics of an "ideal type" TIP evaluation, in this section I focus on what remains to be done and point out opportunities for future research. I also highlight some points not discussed above regarding the challenges of translating this to TIP evaluation.

The first issue that remains to be addressed is related to the evaluation of *policy mixes*. The evaluation of policy mixes is not considered in the integrated framework for TIP evaluation in Paper III, whose focus is program evaluation. However, I would argue that a few insights from the framework could still be harnessed for the evaluation of policy mixes. For example, if we perform the evaluation of multiple programs from a TIP perspective and aggregate the findings using meta-evaluation or secondary analysis (Arnold, 2004; Edler et al., 2008), as discussed in Section 1.1.2, we might be able to capture the systemic effect of policy mixes. Nonetheless, this has not been considered in the scope of this thesis and can serve as a potential avenue for further research. Such efforts could potentially shed light on common causal mechanisms. Still, this should take into consideration contextual aspects. As pointed out by Beach and Pedersen (2019), generalizing the findings from one case to others needs deeper consideration, not only when considering contextual aspects, but also regarding whether we can expect similar processes (or mechanisms) to operate in these other cases. The authors specifically provide guidance on how process tracing can be used to generalize findings through nesting across multiple cases. Other approaches such as qualitative

comparative analysis (QCA) or the Q-method could also be used to clarify and generalize the causal mechanisms (Trein et al., 2021).

QCA can be used to identify causal paths leading to an outcome (Rihoux et al., 2011). For example, it can be used to identify the combination of conditions resulting to certain outcome, based on a number of cases (e.g., projects), and also show different combinations that can also lead to the same outcome.[37] Q-methodology is an approach used to systematically study subjectivity and the "points of view" around a specific topic, using mixed methods (Brown, 1996).[38] It can be used, for example, to analyze the different participants' perspectives regarding the factors that led to the achievement of an outcome as part of the evaluation. Both QCA and Q-methodology, however, would need to be complemented with other approaches, as these would possibly be used to investigate only specific effects of the policy mix.

There have been a few attempts in the TIP literature at policy mix evaluation. For instance, Kivimaa et al. (2017) use a client-oriented evaluation approach to analyze how national policies targeting low-carbon building in Finland contribute to "creative destruction." They use Kivimaa and Kern's (2016) processes to seek the influences of policies in facilitating the zero-carbon transition. Magro and Wilson (2019) address the governance of policy mix evaluation by developing a conceptual framework to analyze the interaction between governance processes in the evaluation policy mix of smart specialization strategies. Purkus and Lüdtke (2020) combine a learning-oriented, systemic approach based on innovation systems and participatory evaluation to propose an evaluation concept for the German instrument mix related to the forest-based bioeconomy governance process. Other authors touch upon the analysis of policy mixes, rather than directly considering their evaluation. For example, Scordato et al. (2018) analyze how the elements of the policy mix supporting the transition of the pulp and paper industry in Sweden interact using the processes proposed by Kivimaa and Kern (2016). Nevertheless, the evaluation

[37] QCA is an approach and a set of techniques developed by Ragin in 1987 that constitutes a middle way between case-oriented and variable-oriented approaches (Rihoux et al., 2011). It aims to gain in-depth insight into cases while also allowing for generalization and can be used to attribute conjunctural causality (Befani, 2013) by identifying the causal paths leading to an outcome, implying that the outcome: (i) can be a result of a combination of conditions ("explanatory variables"), (ii) can be achieved by different combinations of conditions, and (iii) can be impacted differently by a given condition depending on the context (Rihoux et al., 2011).

[38] The approach involves identifying a range of statements through interviews, surveys, or existing literature, which are then classified by participants based on whether they agree, disagree or are neutral in relation to each statement (Herrington and Coogan, 2011). This classification is then factor analyzed, quantitatively, to identify similarities of views between participants on a given issue.

of policy mixes is still underdeveloped in the TIP literature, especially for addressing directionality and additionality, which deserves more attention in future research. Moreover, most of the aforementioned papers also seem to look at instrument mixes rather than the broader definition of policy mixes, as discussed in Section 2.1.

Considering coordination in the evaluation of policy mixes is also an issue that remains unclear in the TIP literature. According to Trein et al. (2021), while the policy literature has largely recognized the importance of coordination, only a few scholars have addressed policy evaluation. This is because there is no clear approach to measure the success and failure of coordination, given the many interactions among different policies, governance levels, and sectors that make it difficult to uncover the effect of policies. Nonetheless, Trein et al. (2021) offer a few suggestions. For example, one suggestion includes advancing empirical analysis about the processes of coordination through systematic reviews or meta-analysis to map the key drivers and effects of coordinated efforts across policy sectors and levels. Another reflects on establishing causality, in which the authors argue that the literature fails to synthesize and combine theoretical and conceptual insights of causal explanations about policy coordination identified in case studies. This can also be done using approaches such as process tracing, QCA, or the Q-method.

Coordination has also been largely overlooked in the TIP literature. A few exceptions include Magro and Wilson (2019), whose conceptual framework considers the coordination of actors, governance, and policy levels. Additionally, Scordato et al. (2018) consider the coordination of policies, governance levels, and sectors when analyzing how policy instruments drive sustainability transitions. However, their approaches do not look at causality or directionality. Therefore, research on coordination in TIP displays gaps in evaluation and offers many opportunities for future research.

The second remaining issue concerns the *broader involvement of stakeholders* in evaluation. Given that TIP involves a broader set of stakeholders, both the impact of different actors on TIP interventions and their involvement in the evaluation itself should be accounted for in TIP evaluation. While the TIP theoretical literature has acknowledged these aspects, they have only been addressed to a limited extent in the literature. Janssen (2019) considers the impact of private firms and other stakeholders in strengthening the TIS functions in his framework. Purkus and Lüdtke's (2020) evaluation concept account for participatory evaluation and reflect how it has

been embedded in their empirical case. The authors also point out some of the main challenges and risks of having a more participatory approach in evaluation, given that the involvement of multiple actors, beyond primary users (e.g. steering and working group members), can deviate the focus of the evaluation due to multiple actors' interests. Kivimaa et al. (2017) take a client-oriented perspective when evaluating of Finnish national policy mixes for low-carbon building. Additionally, Boni et al. (2019) consider participatory evaluation one of the criteria for TIP evaluation and reflect upon key evaluation techniques that account for stakeholder involvement. These include, for instance, multi-criteria mapping (MCM), which involves stakeholders in identifying options and strategies to address a specific goal, and participatory impact pathways analysis (PIPA), which uses participatory workshops to develop program theory.

In the evaluation literature, stakeholder participation in evaluation has been widely accepted for many years, and a few approaches have been put forward (Whitmore, 1999). These include, for example, practical participatory evaluation and transformative participatory evaluation. The former focuses on the evaluation use, in which stakeholders are involved to promote the relevance and ownership of the evaluation, with the goal to support organizational and program decision-making (Cousins and Whitmore, 1998). The latter targets the "democratisation of social change" and, hence, promotes the involvement of various stakeholders, who are empowered to provide their knowledge about the intervention design and implementation. These approaches, however, have mainly been developed within the social justice literature. Therefore, a potential opportunity for future research is to explore participatory evaluation in TIP further using, for example, insights from the evaluation literature on participatory evaluation. Furthermore, the agency of actors in supporting transformative change seem to be largely unaddressed in the TIP literature and also serve as a potential avenue for further research.

Concerning the translational challenges for TIP evaluation introduced in Section 6.3, I have already discussed the aspects related to the issue of accounting for directionality and the effects of TIP in the previous two sections. Nonetheless, some issues remain regarding *developing transformative ToC* and *securing capacity in TIP evaluation*. The former can be considered a third remaining issue in TIP evaluation. The TIP literature, however, does not provide much guidance on how to address such challenges. In Paper II, we discussed some work that can provide insights into developing a more transformative ToC, including the transformative ToC handbook

developed by TIPC (Boni et al., 2021), which uses the transformative outcomes proposed by Ghosh et al. (2021) as the basis for developing the ToC. The evaluation literature also discusses the development of more complex ToC, such as Mayne's (2015) work on useful theories of change and the suggestions of Funnell and Rogers (2011) regarding how to represent complex interventions.

The fourth remaining issue, thus, is securing capacity in TIP evaluation. As pointed out in Paper I, the papers considered in the literature review acknowledged the challenge related to developing institutional and governance capacity in TIP, but did not offer guidance on how to address this. Previous papers have focused on capacity in designing and implementing innovation policies (Borrás, 2011; Borrás and Edquist, 2015). In the context of developing more systemic evaluations, Borrás and Laatsit (2019) acknowledge the need to build capacity at the national level, which would involve identifying the tools that could support countries in taking that step. This might mean that for TIP, building capacity in evaluation needs to be tailored to each country's context. Nonetheless, securing capacity in TIP evaluation might involve education, training, and skills development (i.e., competence building) in different aspects of TIP policymaking, such as sustainability transitions and, specific to evaluation, theory-based approaches and alternative views of causality. Unpacking how to support competence building is therefore another interesting area for further research.

# 8 Conclusion

In Section 7, I discussed my contributions toward addressing the key characteristics of an "ideal type" TIP evaluation, reflected in the feasibility of the integrated framework for TIP program evaluation, as well as highlighting the remaining issues for TIP evaluation. In this section I present the main findings of the thesis and highlight my main contributions to the theory and practice of TIP evaluation.

## 8.1 Main findings

This thesis addressed three main research questions: (RQ1) From a theoretical perspective, what are the distinguishing characteristics of an "ideal type" TIP evaluation compared with previous innovation policy evaluation approaches? (RQ2) How can additionality and directionality be accounted for in TIP program evaluation? (RQ3) What are the challenges related to translating an "ideal type" TIP evaluation approach to practice? To address these research questions, I draw on the five papers appended to this thesis.

Regarding the first research question, I outline five distinguishing characteristics of an "ideal type" TIP evaluation, building mainly on the findings of Papers I and II. The first characteristic includes an evaluation strategy covering different timings, that is, ex-ante, monitoring, interim, and ex-post evaluation, focusing on both formative and developmental evaluation. This is followed by the need to consider entire policy mixes in evaluations. Furthermore, TIP evaluation should also consider the broader involvement of stakeholders. Moreover, TIP evaluation should examine the effects of transformative policies in supporting sustainability transitions and consider directionality. These two last characteristics in particular are described further in Papers III and IV.

Concerning the second research question, I propose an integrated framework for evaluating TIP programs that takes additionality and directionality into consideration. This is mainly covered in Papers III and IV. Paper III presents the integrated framework itself, which has three main components: program theory, system analysis, and synthesis and overall assessment. Paper III also provides a list of steps evaluators can follow when performing a program evaluation. Moreover, it proposes that additionality should go beyond input, output, and behavioral additionality to cover how interventions contribute to the emergence of transformative processes. Thus, the paper

suggests that evaluators should look at "transformative" or system-level additionality. Paper IV describes transformative processes, summarizing the key innovation and transition processes that should underlie TIP program evaluation. Paper IV is one approach among several proposed in the sustainability transition literature to explain the mechanisms and outcomes underlying socio-technical system change. Paper V also contributes to addressing RQ2 by simplifying the integrated framework into five analytical steps and examining the main challenges evaluators face in practice when seeking evidence of additionality and directionality.

Finally, the third research question is addressed mostly by Papers II and V. Both papers identify challenges for the practice of TIP evaluation. Paper II looks at the evaluation practice of three innovation programs implemented by Vinnova: Vinnväxt, CDI, and the SIPs, and Paper V looks specifically at one of the SIPs, namely BioInnovation. Building on insights from these two papers, I then identify four main challenges for TIP evaluation practice. These include accounting for directionality, developing transformative ToC, accounting for the effects of TIP, and securing capacity in TIP evaluation.

## 8.2 Contributions

This thesis offers methodological, conceptual, and empirical contributions. Notably, in terms of its methodological contribution, this thesis constitutes one of the first attempts to address directionality and additionality at the system level in TIP program evaluation through an integrated framework for evaluating TIP programs. To the best of my knowledge, no other framework has considered in detail how to address directionality and additionality at the system level. In relation to this latter aspect, in this thesis, I also advocate a new type of additionality beyond input, output, and behavioral additionality, namely "transformative" or system-level additionality. This is new, as I argue that what has been put forward in previous framings of STI policy evaluation are insufficient to cover the type of additionalities evaluators should be looking for in relation to TIP. Therefore, the integrated framework of Papers III and V reflects an evaluation approach to address this type of additionality based on the notion of transformative processes proposed in Paper IV.

Conceptually, the thesis also summarizes the key characteristic of an "ideal type" TIP evaluation, which can indicate the most important points to be considered by evaluators when assessing TIP.

This is summarized based on insights from the systematic literature review of TIP covered in Paper I and the analytical framework characterizing an "ideal type" TIP in Paper II.

While the integrated framework from Paper III provides mainly methodological contributions, indirectly, it also includes conceptual work when it seeks to combine insights from the TIP literature, which refers to the social science theory underlying the evaluation, and the theory-based approach to evaluation. Moreover, the set of transformative outcomes from Paper IV contributes conceptually to widening the TIS function framework to include actor networks and institutional change processes that were not captured previously, while also accounting for overlaps in the literature that could be linked to already established functions. This adds to the understanding of what kind of transformative outcomes the evaluator can look for when assessing the impact of TIP intervention.

Empirically, the analysis of the current practice of evaluation in three Swedish programs (Vinnväxt, CID, and SIPs) contributes to the practical understanding of the main challenges innovation agencies such as Vinnova might need to consider when aligning its evaluation strategy with a TIP orientation. Additionally, I offer insights through Paper V on the main results of the examination of an ongoing SIP, that is, BioInnovation, which constitutes a case study of how an evaluation of TIP could look like.

## 8.3 Limitations and recommendations for future research

The limitations and recommendations for future research have been already addressed in Sections 7.2 and 7.3. In sum, I reflect on the main limitations of the integrated framework and the feasibility of applying it in practice. First, I pointed out that it can be difficult (but not impossible) to trace causal mechanisms in TIP programs to shed light on additionality. This is especially affected by the limitations of the data, which, in the case of BioInnovation presented in Paper V, were insufficient to provide much evidence for the confirmation of causal mechanisms. Second, I also highlighted that identifying the overall trajectory of change to assess directionality might not be feasible during a program's lifespan. Although in the BioInnovation case we could observe some patterns emerging related to the transition pathways, they might not be enough to reveal whether a transition is unfolding due to the long-term character of socio-technical change. Finally, I

recognized that one empirical application of the framework is still not enough to show its entire potential and further refine it to be more practice-oriented.

For future research, I suggest that more empirical applications of the integrated framework should be developed, for example, to other SIPs or even to cases outside Sweden. This could also involve testing other approaches to assessing directionality. Moreover, a broadened concept of system-level additionality requires further research to reflect on its implications for both theory and practice.

A few recommendations can also be drawn based on the TIP distinguishing characteristics and remaining challenges not addressed in this thesis. For example, in relation to the evaluation of policy mixes, I suggested using secondary analysis or meta-evaluation based on the application of the framework to different programs. This can serve as a basis for the identification and generalization of causal mechanisms, for example, by applying approaches such as process tracing, QCA, and the Q-method. Additionally, in relation to coordination, I pointed out that this has been understudied in both the policy evaluation and TIP literature. Therefore, suggestions for further research include advancing the empirical analysis of the process of coordination through systematic reviews or meta-analyses or by clarifying and generalizing causal mechanisms regarding policy coordination (Trein et al., 2021). This latter suggestion specifically can be based on the findings of case studies, and approaches such as process tracing, QCA or Q-method can support the investigation of causal relations.

Finally, concerning the need to involve stakeholders in evaluation more broadly, I recommended that approaches suggested in the evaluation literature, such as practical participatory evaluation and transformative participatory evaluation, could potentially be harnessed for TIP evaluation. Furthermore, I pointed out that the agency of actors in supporting transformative change have been unaddressed in the TIP literature and can also serve as a potential avenue for further research.

Two remaining issues related to the practical challenges in the translation of TIP theory are developing more transformative ToCs and securing capacity in TIP evaluation. I discuss a few works that can shed light on the former issue. However, the second issue remains largely unaddressed in the literature and deserves further attention.

# References


Alkin, M.C., 2013. Evaluation Roots: A Wider Perspective of Theorists' Views and Influences, 2nd ed. SAGE Publications, Thousand Oaks, US.

Amanatidou, E., Cunningham, P., Gök, A., Garefi, I., 2014. Using evaluation research as a means for policy analysis in a 'new' mission-oriented policy context. Minerva 52, 419–438.

Andersson, J., 2020. Shape it until you make it: a conceptual foundation for efforts to analyze and shape technological innovation Technology Management and Economics. PhD thesis. Chalmers University of Technology, Gothenburg.

Andersson, J., Hellsmark, H., Sandén, B., 2021. The outcomes of directionality: Towards a morphology of sociotechnical systems. Environmental Innovation and Societal Transitions 40, 108–131.

Arnold, E., 2004. Evaluating research and innovation policy: A systems world needs systems evaluations. Research Evaluation 13, 3–17.

Arnold, E., Åström, T., Glass, C., De Scalzi, M., 2018. How should we evaluate complex programmes for innovation and socio-technical transitions? (downloaded on 23 February 2021 from https://www.technopolis-group.com/wp-content/uploads/2020/02/How-should-we-evaluate-complex-programmes-for-innovation-and-socio-technical-transitions.pdf).

Arrow, K., 1962. Economic welfare and the allocation of resources for invention, in: Nelson, R. (Ed.), The Rate and Direction of Inventive Activity. Princeton University Press, Princeton, pp. 609–625.

Astbury, B., 2013. Some reflections on Pawson's science of evaluation: A realist manifesto. Evaluation 19, 383–401.

Astbury, B., Leeuw, F.L., 2010. Unpacking black boxes: Mechanisms and theory building in evaluation. American Journal of Evaluation 31, 363–381.

Autio, E., Kanninen, S., Gustafsson, R., 2008. First-and second-order additionality and learning outcomes in collaborative R&D programs. Research Policy 37, 59–76.

Bach, L., Matt, M., 2005. From economic foundations to S&T policy tools: A comparative analysis of the dominant paradigms, in: Llerena, P., Matt, M. (Eds.), Innovation Policy in a Knowledge-Based Economy: Theory and Practice. Springer Berlin Heidelberg, Berlin, pp. 17–45.

Beach, D., Pedersen, R.B., 2019. Process-tracing methods: foundations and guidelines, 2nd ed. University of Michigan Press, Ann Arbor.

Befani, B., 2013. Between complexity and generalization: Addressing evaluation challenges with QCA. Evaluation 19, 269–283.

Bellandi, M., Caloffi, A., 2010. An Analysis of Regional Policies Promoting Networks for Innovation. European Planning Studies 18, 67-82.

Bergek, A., 2004. Elcertifikatsystemet och innovationssystemen inom förnybar energiteknik. (The Electricity Certificate System and the Innovation Systems for Renewable Energy Technology) (downloaded on 19 April 2021 from https://publications.lib.chalmers.se/records/fulltext/218305/local_218305.pdf).

Bergek, A., 2019. Technological Innovation Systems: A review of recent findings and suggestions for future research, in: Boons, F., McMeekin, A. (Eds.), Handbook of Sustainable Innovation. Edvard Elgar Publishing, Cheltenham, pp. 200–218.

Bergek, A., Bjørgum, Ø., Hansen, T., Hanson, J., Steen, M., 2021. Sustainability transitions in coastal shipping: The role of regime segmentation. Transportation Research Interdisciplinary Perspectives 12, 100497.

Bergek, A., Haddad, C.R., 2022. Evaluating transformative innovation policy outcomes as unfolding processes of change in sociotechnical configurations, in: Vélez-Cuartas, G., Romero-Goyeneche, O.Y. (Eds.), Transformative Metrics: Contributions to the Studies for Monitoring and Evaluating How Science, Technology, and Innovation Can Address Social and Environmental Callenges. Universidad de Antioquia, Medellín.

Bergek, A., Hekkert, M., Jacobsson, S., Markard, J., Sandén, B., Truffer, B., 2015. Technological innovation systems in contexts: Conceptualizing contextual structures and interaction dynamics. Environmental Innovation and Societal Transitions 16, 51–64.

Bergek, A., Hellsmark, H., Karltorp, K., 2023. Directionality challenges for transformative innovation policy: Lessons from implementing climate goals in the process industry. Industry and Innovation, 1–30.

Bergek, A., Jacobsson, S., Carlsson, B., Lindmark, S., Rickne, A., 2008. Analyzing the functional dynamics of technological innovation systems: A scheme of analysis. Research Policy 37, 407–429.

Bergek, A., Jacobsson, S., Hekkert, M., Smith, K., 2010. Functionality of innovation systems as a rationale and guide in innovation policy, in: Smits, R., Kuhlmann, S., Shapira, P. (Eds.), The Theory and Practice of Innovation Policy. Edward Elgar, Cheltenham, pp. 117–146.

Bijker, W.E., Hughes, T.P., Trevor, J., 1984. Pinch, eds. 1987. The social construction of technological systems: New directions in the sociology and history of technology. MIT Press, Cambridge.

BioInnovation, 2020a. The circular bioeconomy shapes the society of the future (downloaded on 19 March 2021 from https://www.e-magin.se/paper/d5609c8r/paper/1#/paper/d5609c8r/1).

BioInnovation, 2020b. The circular bioeconomy shapes the society of the future: research and innovation for sustainable development (accessed on 11 December 2020 from https://www.bioinnovation.se/en/agenda/).

BioInnovation, 2023a. Om oss (accessed on 30 June 2023 from https://www.bioinnovation.se/).

BioInnovation, 2023b. Projekt (Project) (accessed on 21 March 2023 from https://www.bioinnovation.se/projekt/).

Blamey, A., Mackenzie, M., 2007. Theories of change and realistic evaluation: Peas in a pod or apples and oranges? Evaluation 13, 439–455.

Boni, A., Giachi, S., Molas-Gallart, J., 2019. Towards a framework for transformative innovation policy evaluation (downloaded on 16 February 2023 from https://www.tipconsortium.net/wp-content/uploads/2019/07/Towards-a-Framework-for-TIP-Evaluation-Designed.pdf).

Boni, A., Penna, C.C.R., Palavitino, C.A., Brodnik, C., Matti, C., Witte, J., Schot, J., Molas-Gallart, J., Weber, M., Mendez, P.F., Terrazas, P., Keesman, S., 2021. Motion handbook: Developing a transformative theory of change (downloaded on 15 March 2023 from https://www.tipconsortium.net/publication/motion-handbook-developing-a-transformative-theory-of-change/).

Borrás, S., 2011. Policy learning and organizational capacities in innovation policies. Science and Public Policy 38, 725–734.

Borrás, S., Edquist, C., 2013. The choice of innovation policy instruments. Technological Forecasting and Social Change 80, 1513–1522.

Borrás, S., Edquist, C., 2015. Education, training and skills in innovation policy. Science and Public Policy 42, 215–227.

Borrás, S., Laatsit, M., 2019. Towards system oriented innovation policy evaluation? Evidence from EU28 member states. Research Policy 48, 312–321.

Borrás, S., Schwaag Serger, S., 2022. The design of transformative research and innovation policy instruments for grand challenges: The policy-nesting perspective. Science and Public Policy 49, 659–672.

Boschma, R., 2005. Proximity and Innovation: A Critical Assessment. Regional Studies 39, 61-74.

Bowen, G.A., 2009. Document analysis as a qualitative research method. Qualitative research journal 9, 27–40.

Brown, S.R., 1996. Q methodology and qualitative research. Qualitative health research 6, 561–567.

Bryman, A., Bell, E., 2015. Business research methods, 4th ed. Oxford University Press, Oxford.

Bugge, M., Coenen, L., Marques, P., Morgan, K., 2017a. Governing system innovation: assisted living experiments in the UK and Norway. European Planning Studies 25, 2138-2156.

Bugge, M., Coenen, L., Marques, P., Morgan, K., 2017b. Governing system innovation: assisted living experiments in the UK and Norway. European Planning Studies 25, 2138–2156.

Bugge, M.M., Coenen, L., Branstad, A., 2018. Governing socio-technical change: Orchestrating demand for assisted living in ageing societies. Science and Public Policy 45, 468–479.

Cairney, P., 2012. What is public policy? How should we study it?, in: Cairney, P. (Ed.), Understanding Public Policy: Theories and Issues. Palgrave Macmillan, London, pp. 22–45.

Campbell, D.S.J., 1966. Experimental and quasi-experimental designs for research. Rand McNally, Chicago.

Cappellano, F., Makkonen, T., Dotti, N.F., Morisson, A., Rizzo, A., 2022. Where innovation meets directionality: An index to measure regional readiness to deal with societal challenges. European Planning Studies 30, 1549–1576.

Carlsson, B., 1995. Technological Systems and Economic Performance: the case of factory automation, in: Dodgson, M., Rothwell, R. (Eds.), The Handbook of Industrial Innovation. Edward Elgar Publishing, Cheltenham, pp. 13-24.

Carlsson, B., Jacobsson, S., 1997. In search of useful public policies - key lessons and issues for policy makers, in: Carlsson, B. (Ed.), Technological Systems and Industrial Dynamics. Kluwer Academic Publishers, Dordrecht, pp. 299-315.

Carlsson, B., Jacobsson, S., Holmén, M., Rickne, A., 2002. Innovation systems: Analytical and methodological issues. Research Policy 31, 233–245.

Carlsson, B., Stankiewicz, R., 1991. On the nature, function and composition of technological systems. Journal of Evolutionary Economics 1, 93–118.

Cevallos, R.A., Merino Moreno, C., 2022. Collegial forms of implementation of directionality in national innovation strategies. Foresight and STI Governance 16, 46–58.

Chaminade, C., Edquist, C., 2008. Rationales for public policy intervention in the innovation process: A systems of innovation approach, in: Smits, R.E., Kulhmann, S., Shapira, P. (Eds.), The Theory and Practice of Innovation Policy. Edward Elgar, Cheltenham, pp. 97-116.

Chaminade, C., Edquist, C., 2010. Rationales for public policy intervention in the innovation process: A systems of innovation approach, in: Smits, R.E., Kuhlmann, S., Shapira, P. (Eds.), The Theory and Practice of Innovation Policy: An International Research Handbook. Edward Elgar Publishing, Cheltenham, pp. 95-114.

Chen, H.-T., 1996. A comprehensive typology for program evaluation. Evaluation practice 17, 121-130.

Chen, H.T., 1990. Theory-driven evaluations. SAGE Publications, Newbury Park.

Christie, C.A., Fleischer, D., 2009. Social inquiry paradigms as a frame for the debate on credible evidence, in: Donaldson, S.I., Christie, C.A., Mark, M.M. (Eds.), What counts as credible evidence in applied research and evaluation practice. SAGE Publications, Thousand Oaks, pp. 19-30.

Clarysse, B., Wright, M., Mustar, P., 2009. Behavioural additionality of R&D subsidies: A learning perspective. Research Policy 38, 1517–1533.

Connell, J.P., Kubisch, A.C., Schorr, L.B., Weiss, C.H., 1995. New Approaches to Evaluating Community Initiatives. Aspen Institute, Washington.

Cooke, P., Uranga, M.G., Etxebarria, G., 1997. Regional innovation systems: Institutional and organisational dimensions. Research Policy 26, 475–491.

Cousins, J.B., Whitmore, E., 1998. Framing participatory evaluation. New directions for evaluation 1998, 5-23.

Cronbach, L.J., Shapiro, K., 1982. Designing evaluations of educational and social programs. Jossey-Bass, San Francisco.

de Oliveira, L.G.S., Lacerda, J.S., Negro, S.O., 2020. A mechanism-based explanation for blocking mechanisms in technological innovation systems. Environmental Innovation and Societal Transitions 37, 18–38.

de Oliveira, L.G.S., Negro, S.O., 2019. Contextual structures and interaction dynamics in the Brazilian Biogas Innovation System. Renewable and Sustainable Energy Reviews 107, 462–481.

Diercks, G., 2018. Lost in translation: How legacy limits the OECD in promoting new policy mixes for sustainability transitions. Research Policy 48, 103667.

Diercks, G., Larsen, H., Steward, F., 2019. Transformative innovation policy: Addressing variety in an emerging policy paradigm. Research Policy 48, 880–894.

Donaldson, S.I., Lipsey, M.W., 2006. Roles for theory in contemporary evaluation practice: Developing practical knowledge, in: Shaw, I.F., Greene, J.C., Mark , M.M. (Eds.), The handbook of evaluation: Policies, programs, and practices. SAGE Publications, Thousand Oaks, pp. 56-75.

Edelenbos, J., Van Buuren, A., 2005. The learning evaluation: a theoretical and empirical exploration. Evaluation review 29, 591-612.

Edler, J., Berger, M., Dinges, M., Gök, A., 2012. The practice of evaluation in innovation policy in Europe. Research Evaluation 21, 167-182.

Edler, J., Ebersberger, B., Lo, V., 2008. Improving policy understanding by means of secondary analyses of policy evaluation. Research Evaluation 17, 175–186.

Edler, J., Fagerberg, J., 2017. Innovation policy: What, why, and how. Oxford Review of Economic Policy 33, 2–23.

Eisenhardt, M.E., 1989. Building Theories from Case Study Research. The Academy of Management Review 14, 532-550.

Elzinga, R., Janssen, M.J., Wesseling, J., Negro, S.O., Hekkert, M.P., 2023. Assessing mission-specific innovation systems: Towards an analytical framework. Environmental Innovation and Societal Transitions 48, 100745.

Etzkowitz, H., Leydesdorff, L., 2000. The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university–industry–government relations. Research Policy 29, 109-123.

Fagerberg, J., 2017. Innovation policy: rationales, lessons and challenges. Journal of Economic Surveys 31, 497–512.

Flanagan, K., Uyarra, E., Laranja, M., 2011. Reconceptualising the 'policy mix' for innovation. Research Policy 40, 702–713.

Freeman, C., 1987. Technology Policy and Economic Performance: Lessons from Japan Pinter, London.

Fulbright-Anderson, K., Connell, J.P., Kubisch, A.C., 1998. New approaches to evaluating community initiatives: Theory, measurement, and analysis. Aspen Institute, Washington.

Funnell, S.C., Rogers, P.J., 2011. Purposeful Program Theory: Effective Use of Theories of Change and Logic Models. John Wiley & Sons San Francisco.

Gates, E., Dyson, L., 2017. Implications of the changing conversation about causality for evaluators. American Journal of Evaluation 38, 29-46.

Geels, F., Raven, R., 2006. Non-linearity and Expectations in Niche-Development Trajectories: Ups and Downs in Dutch Biogas Development (1973–2003). Technology Analysis & Strategic Management 18, 375-392.

Geels, F.W., 2002. Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. Research Policy 31, 1257-1274.

Geels, F.W., 2004. From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. Research Policy 33, 897-920.

Geels, F.W., 2022. Causality and explanation in socio-technical transitions research: Mobilising epistemological insights from the wider social sciences. Research Policy 51.

Geels, F.W., Kern, F., Fuchs, G., Hinderer, N., Kungl, G., Mylan, J., Neukirch, M., Wassermann, S., 2016. The enactment of socio-technical transition pathways: A reformulated typology and a comparative multi-level analysis of the German and UK low-carbon electricity transitions (1990–2014). Research Policy 45, 896–913.

Geels, F.W., Schot, J., 2007. Typology of sociotechnical transition pathways. Research Policy 36, 399–417.

Georghiou, L., 1994. Impact of the Framework Programme on European Industry (downloaded on 28 June 2021 from https://op.europa.eu/en/publication-detail/-/publication/f4592590-8dc0-4012-afa6-1482119a7024).

Georghiou, L., 1998. Issues in the Evaluation of Innovation and Technology Policy. Evaluation 4, 37–51.

Georghiou, L., Clarysse, B., 2006. Government R&D funding and company behaviour: measuring behavioural additionality (downloaded on 17 June 2020 from https://www.oecd-ilibrary.org/science-and-technology/government-r-d-funding-and-company-behaviour_9789264025851-en).

Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B., Vermeersch, C.M.J., 2011. Impact Evaluation in Practice The World Bank, Washington.

Ghosh, B., Kivimaa, P., Ramirez, M., Schot, J., Torrens, J., 2021. Transformative Outcomes: Assessing and reorienting experimentation with transformative innovation policy. Science and Public Policy 48, 739–756.

Grillitsch, M., Hansen, T., Coenen, L., Miörner, J., Moodysson, J., 2019. Innovation policy for system-wide transformation: The case of strategic innovation programmes (SIPs) in Sweden. Research Policy 48, 1048-1061.

Guba, E.G., Lincoln, Y.S., 1989. Fourth Generation Evaluation. SAGE Publications, Thousand Oaks.

Gök, A., 2010. Evolutionary Approach to Innovation Policy Evaluation: Behavioural Additionality and Organisational Routines, Faculty of Humanities. University of Manchester Manchester, p. 289.

Gök, A., Edler, J., 2012. The use of behavioural additionality evaluation in innovation policy making. Research Evaluation 21, 306–318.

Haddad, C.R., 2021a. Evaluating transformative innovation policy: A systems analysis of the BioInnovation Strategic Innovation Program in Sweden, 12th International Sustainability Transitions Conference, Karlsruhe

Haddad, C.R., 2021b. Evaluating transformative innovation policy: Towards an integrated framework, Department of Technology Management and Economics. Licentiate Thesis. Chalmers University of Technology, Gothenburg.

Haddad, C.R., Bergek, A., 2020. A functions approach for evaluating transformative innovation policy, 11th International Sustainability Transition Conference, 18-21 August, Vienna, Austria.

Haddad, C.R., Bergek, A., 2023. Towards an integrated framework for evaluating transformative innovation policy. Research Policy 52, 104676.

Haddad, C.R., Nakić, V., Bergek, A., Hellsmark, H., 2022. Transformative innovation policy: A systematic review. Environmental Innovation and Societal Transitions 43, 14–40.

Hall, P.A., 1993. Policy paradigms, social learning, and the state: the case of economic policymaking in Britain. Comparative politics, 275-296.

Hekkert, M.P., Janssen, M.J., Wesseling, J.H., Negro, S.O., 2020. Mission-oriented innovation systems. Environmental Innovation and Societal Transitions 34, 76–79.

Hekkert, M.P., Suurs, R.A., Negro, S.O., Kuhlmann, S., Smits, R.E., 2007. Functions of innovation systems: A new approach for analysing technological change. Technological Forecasting and Social Change 74, 413-432.

Herrington, N., Coogan, J., 2011. Q methodology: an overview. Research in Teacher Education 1, 24-28.

Hind, J., 2010. Additionality: A useful way to construct the counterfactual qualitatively? Evaluation Journal of Australasia 10, 28–35.

Hoogma, R., 2000. Exploiting Technological Niches: Strategies for Experimental Introduction of Electric Vehicles. PhD thesis. Twente University of Technology, Enschede.

Hoppmann, J., Huenteler, J., Girod, B., 2014. Compulsive policy-making—The evolution of the German feed-in tariff system for solar photovoltaic power. Research Policy 43, 1422–1441.

House, E.R., 2001. Unfinished business: Causes and values. American Journal of Evaluation 22, 309-315.

Howlett, M., Cashore, B., 2020. Public policy: definitions and approaches, in: Cashore, B., Howlett, M. (Eds.), A modern guide to public policy. Edward Elgar Publishing Ltd., Cheltenham, UK, pp. 10-21.

Howlett, M., Giest, S., 2013. The policy-making process, in: Araral Jr., E., Fritzen, S., Howlett, M., Ramesh, M., Wu, X. (Eds.), Routledge Handbook of Public Policy. Routledge, New York.

Howlett, M., Rayner, J., 2007. Design principles for policy mixes: Cohesion and coherence in 'new governance arrangements'. Policy and Society 26, 118.

Hölscher, K., Wittmayer, J.M., Loorbach, D., 2018. Transition versus transformation: What's the difference? Environmental Innovation and Societal Transitions 27, 1–3.

Isaksen, A., 1999. Evaluation of a regional innovation programme: the Innovation and New Technology Programme in Northern Norway. Evaluation and program planning 22, 83–90.

Jaakkola, E., 2020. Designing conceptual articles: four approaches. AMS review 10, 18-26.

Jacobsson, S., Bergek, A., 2004. Transforming the energy sector: the evolution of technological systems in renewable energy technology. Industrial and Corporate Change 13, 815-849.

Jacobsson, S., Bergek, A., Sandén, B., 2017. Improving the European Commission's analytical base for designing instrument mixes in the energy sector: Market failures versus system weaknesses. Energy Research & Social Science 33, 11-20.

Jacobsson, S., Johnson, A., 2000. The diffusion of renewable energy technology: An analytical framework and key Issues for research. Energy Policy 28, 625–640.

Jacobsson, S., Perez Vico, E., 2010. Towards a systemic framework for capturing and explaining the effects of academic R&D. Technology Analysis & Strategic Management 22, 765-787.

Janssen, M.J., 2019. What bangs for your buck? Assessing the design and impact of Dutch transformative policy. Technological Forecasting and Social Change 138, 78–94.

Janssen, M.J., Bergek, A., Wesseling, J.H., 2022. Evaluating systemic innovation and transition programmes: Towards a culture of learning. PLOS Sustainability and Transformation 1, e0000008.

Johnson, A., Jacobsson, S., 2001. Inducement and blocking mechanisms in the development of a new industry: the case of renewable energy technology in Sweden, in: Coombs, R., Green, K., Richards, A., Walsh, V. (Eds.), Technology and the market: demand, users and innovation. Edward Elgar Publishing Cheltenham, pp. 89-111.

Kemp, R., 2010. Eco-innovation: definition, measurement and open research issues. Economia politica 27, 397–420.

Kemp, R., Loorbach, D., 2006. Transition management: a reflexive governance approach, in: Voß, J.-P., Bauknecht, D., Kemp, R. (Eds.), Reflexive Governance for Sustainable Development. Edward Elgar, Cheltenham,, pp. 103-130.

Kemp, R., Rotmans, J., 2005. Transition management: managing the co-evolution of technical, environmental and social systems, in: Weber, K.M., Hemmelskamp, J. (Eds.), Towards Environmental Innovation Systems. Springer, Heidelberg, pp. 33–55.

Kemp, R., Schot, J., Hoogma, R., 1998. Regime shifts to sustainability through processes of niche formation: The approach of strategic niche management. Technology Analysis & Strategic Management 10, 175-198.

Kern, F., 2012. Using the multi-level perspective on socio-technical transitions to assess innovation policy. Technological Forecasting and Social Change 79, 298-310.

Kern, F., Rogge, K.S., 2018. Harnessing theories of the policy process for analysing the politics of sustainability transitions: A critical survey. Environmental Innovation and Societal Transitions 27, 102–117.

Khakee, A., 2003. The emerging gap between evaluation research and practice. Evaluation 9, 340–352.

Khandker, S.R., Koolwal, G.B., Samad, H.A., 2010. Handbook on Impact Evaluation: Quantitative Methods and Practices. The World Bank, Washington.

Kivimaa, P., Kangas, H.-L., Lazarevic, D., 2017. Client-oriented evaluation of 'creative destruction' in policy mixes: Finnish policies on building energy efficiency transition. Energy research & social science 33, 115–127.

Kivimaa, P., Kern, F., 2016. Creative destruction or mere niche support? Innovation policy mixes for sustainability transitions. Research Policy 45, 205–217.

Kivimaa, P., Virkamäki, V., 2014. Policy mixes, policy interplay and low carbon transitions: The case of passenger transport in Finland. Environmental Policy and Governance 24, 28–41.

Kline, S.J., Rosenberg, N., 1986. An overview of innovation, in: Landau, R., Rosenberg, N. (Eds.), The Positive Sum Strategy: Harnessing Technology for Economic Growth. National Academy of Sciences, Washington, pp. 275-306.

Knill, C., Tosun, J., 2012. Public Policy: A New Introduction. Palgrave Macmillan, London.

Kubera, P., 2021. The challenges in the evaluation of behavioural additionality of innovation policy. The Journal of Organizational Management Studies 2021, 1–11.

Köhler, J., Geels, F.W., Kern, F., Markard, J., Onsongo, E., Wieczorek, A., Alkemade, F., Avelino, F., Bergek, A., Boons, F., 2019. An agenda for sustainability transitions research: State of the art and future directions. Environmental Innovation and Societal Transitions 31, 1–32.

Larrue, P., 2021. The design and implementation of mission-oriented innovation policies: A new systemic policy approach to address societal challenges (downloaded on 21 February 2023 from https://www.oecd.org/sti/inno/the-design-and-implementation-of-mission-oriented-innovation-policies-3f6c76a4-en.htm).

Leeuw, F., Vaessen, J., 2009. Impact Evaluations and Development: Nonie Guidance on Impact Evaluation (downloaded on from https://documents1.worldbank.org/curated/en/411821468313779505/pdf/574900WP0Box351BLIC10nonie1guidance.pdf).

Lehmann, P., 2012. Justifying a policy mix for pollution control: a review of economic literature. Journal of Economic Surveys 26, 71–97.

Lindberg, M.B., Markard, J., Andersen, A.D., 2019. Policies, actors and sustainability transition pathways: A study of the EU's energy policy mix. Research Policy 48, 103668.

Loorbach, D., Frantzeskaki, N., Avelino, F., 2017. Sustainability transitions research: Transforming science and practice for societal change. Annual Review of Environment and Resources 42, 599–626.

Loorbach, D., Rotmans, J., 2010. The practice of transition management: Examples and lessons from four distinct cases. Futures 42, 237–246.

Lundvall, B.-Å., 1992. National systems of innovation: Towards a theory of innovation and interactive learning. Pinter, London.

Lundvall, B.-Å., Borrás, S., 2005. Science, technology and innovation policy, in: Fagerberg, J., Mowery, D.C., Nelson, R.R. (Eds.), The Oxford handbook of innovation. Oxford University Press, Oxford, pp. 599-631.

Luukkonen, T., 2000. Additionality of EU framework programmes. Research Policy 29, 711–724.

Luukkonen, T., 2002. Research evaluation in Europe: state of the art. Research Evaluation 11, 81-84.

Mackenzie, M., Blamey, A., 2005. The practice and the theory: lessons from the application of a theories of change approach. Evaluation 11, 151-168.

Magnusson, T., Berggren, C., 2018. Competing innovation systems and the need for redeployment in sustainability transitions. Technological Forecasting and Social Change 126, 217-230.

Magro, E., Wilson, J.R., 2013. Complex innovation policy systems: Towards an evaluation mix. Research Policy 42, 1647–1656.

Magro, E., Wilson, J.R., 2019. Policy-mix evaluation: Governance challenges from new place-based innovation policies. Research Policy 48, 103612.

Malerba, F., 2002. Sectoral systems of innovation and production. Research Policy 31, 247-264.

Markard, J., Hekkert, M., Jacobsson, S., 2015. The technological innovation systems framework: Response to six criticisms. Environmental Innovation and Societal Transitions 16, 76–86.

Markard, J., Raven, R., Truffer, B., 2012. Sustainability transitions: An emerging field of research and its prospects. Research Policy 41, 955–967.

Markard, J., Truffer, B., 2008. Technological innovation systems and the multi-level perspective: Towards an integrated framework. Research Policy 37, 596–615.

Marques Santos, A., Coad, A., 2023. Suggestions for Monitoring and Evaluation of Transformative Innovation Policy (downloaded on 02 June 2023 from https://publications.jrc.ec.europa.eu/repository/handle/JRC132492).

Martin, B.R., 2012. The evolution of science policy and innovation studies. Research Policy 41, 1219–1239.

Mathison, S., 2005. Encyclopedia of evaluation. Sage Publications, Thousand Oaks.

Mayne, J., 2015. Useful theory of change models. Canadian Journal of Program Evaluation 30.

Mazzucato, M., 2016. From market fixing to market-creating: A new framework for innovation policy. Industry and Innovation 23, 140–156.

Mazzucato, M., 2018. Mission-oriented innovation policies: Challenges and opportunities. Industrial and Corporate Change 27, 803-815.

Meelen, T., Farla, J., 2013. Towards an integrated framework for analysing sustainable innovation policy. Technology Analysis & Strategic Management 25, 957-970.

Miedzinski, M., Mazzucato, M., Ekins, P., 2019. A framework for mission-oriented innovation policy roadmapping for the SDGs: The case of plastic-free oceans (downloaded on 14 March 2023 from https://apo.org.au/node/252636).

Molas-Gallart, J., Boni, A., Giachi, S., Schot, J., 2021. A formative approach to the evaluation of Transformative Innovation Policies. Research Evaluation 30, 431–442.

Molas-Gallart, J., Davies, A., 2006. Toward theory-led evaluation: The experience of European science, technology, and innovation policies. American Journal of Evaluation 27, 64–82.

Naber, R., Raven, R., Kouw, M., Dassen, T., 2017. Scaling up sustainable energy innovations. Energy Policy 110, 342-354.

Nelson, R.R., 1959. The simple economics of basic scientific research. Journal of Political Economy 67, 297-306.

Nelson, R.R., 1993. National Innovation Systems: A Comparative Analysis. Oxford University Press on Demand, Oxford, UK.

Nelson, R.R., Winter, S.G., 1982. The Schumpeterian tradeoff revisited. The American Economic Review 72, 114-132.

Nelson, R.R., Winter, S.G., 1985. An Evolutionary Theory of Economic Change. Harvard University Press, Cambridge.

OECD, 1992. OECD Proposed Guidelines for Collecting and Interpreting Technological Innovation Data: Oslo Manual. OECD Publishing, Paris.

OECD, 2015. Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development (downloaded on 05 February 2021 from https://www.oecd.org/publications/frascati-manual-2015-9789264239012-en.htm).

OECD, 2016. OECD Reviews of Innovation Policy: Sweden 2016 (downloaded on 15 July 2020 from http://www.oecd.org/publications/oecd-reviews-of-innovation-policy-sweden-2016-9789264250000-en.htm).

Parks, D., 2022. Directionality in transformative innovation policy: Who is giving directions? Environmental Innovation and Societal Transitions 43, 1–13.

Patton, M.Q., 2010. Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use. The Guilford Press, New York.

Pawson, R., 2006. Evidence-based policy: a realist perspective. SAGE Publications, Thousand Oaks.

Pawson, R., 2013. The Science of Evaluation: a Realist Manifesto SAGE Publications, London.

Pawson, R., Greenhalgh, T., Harvey, G., Walshe, K., 2005. Realist review-a new method of systematic review designed for complex policy interventions. Journal of health services research & policy 10, 21-34.

Pawson, R., Tilley, N., 1997. Realist Evaluation. SAGE Publications, London.

Pel, B., Raven, R., van Est, R., 2020. Transitions governance with a sense of direction: synchronization challenges in the case of the Dutch 'Driverless Car' transition. Technological Forecasting and Social Change 160, 120244.

Perez Vico, E., Jacobsson, S., 2012. Identifying, explaining and improving the effects of academic R&D: the case of nanotechnology in Sweden. Science and Public Policy 39, 513-529.

Petticrew, M., Roberts, H., 2008. Systematic Reviews in the Social Sciences: A practical Guide. John Wiley & Sons, New York, US.

Purkus, A., Lüdtke, J., 2020. A systemic evaluation framework for a multi-actor, forest-based bioeconomy governance process: The German Charter for Wood 2.0 as a case study. Forest policy and economics 113, 102113.

Rametsteiner, E., Weiss, G., 2006. Assessing policies from a systems perspecitve—Experiences with applied innovation systems analysis and implications for policy evaluation. Forest policy and economics 8, 564-576.

Raven, R., Kern, F., Verhees, B., Smith, A., 2016. Niche construction and empowerment through socio-political work. A meta-analysis of six low-carbon technology cases. Environmental Innovation and Societal Transitions 18, 164-180.

Rihoux, B., Rezsöhazy, I., Bol, D., 2011. Qualitative Comparative Analysis (QCA) in Public Policy Analysis: an Extensive Review. German Policy Studies/Politikfeldanalyse 7.

Rip, A., Kemp, R., 1998. Technological change, in: Rayner, S., Malone, E.L. (Eds.), Human Choice and Climate Change. Battelle Press, Columbus, pp. 327–399.

Roessner, J.D., 2002. Outcome measurement in the USA: State of the art. Research Evaluation 11, 85–93.

Rogers, P.J., Weiss, C.H., 2007. Theory-based evaluation: Reflections ten years on: Theory-based evaluation: Past, present, and future. New directions for evaluation 114.

Rogge, K.S., Reichardt, K., 2016. Policy mixes for sustainability transitions: An extended concept and framework for analysis. Research Policy 45, 1620–1635.

Rohracher, H., Coenen, L., Kordas, O., 2022. Mission incomplete: Layered practices of monitoring and evaluation in Swedish transformative innovation policy. Science and Public Policy.

Rolfe, S., 2016. Assessing the impacts of community participation policy and practice in Scotland and England, School of Social and Political Sciences. PhD Thesis. University of Glasgow Glasgow, p. 426.

Rolfe, S., 2019. Combining Theories of Change and Realist Evaluation in practice: Lessons from a research on evaluation study. Evaluation 25, 294-316.

Rotmans, J., Kemp, R., Van Asselt, M., 2001. More evolution than revolution: transition management in public policy. Foresight.

Russo, M., Rossi, F., 2009. Cooperation networks and innovation: A complex systems perspective to the analysis and evaluation of a regional innovation policy programme. Evaluation 15, 75-99.

Schlaile, M.P., Urmetzer, S., Blok, V., Andersen, A.D., Timmermans, J., Mueller, M., Fagerberg, J., Pyka, A., 2017. Innovation systems for transformations towards sustainability? Taking the normative dimension seriously. Sustainability 9, 2253.

Schot, J., Geels, F.W., 2008. Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy. Technology Analysis & Strategic Management 20, 537-554.

Schot, J., Steinmueller, W.E., 2018. Three frames for innovation policy: R&D, systems of innovation and transformative change. Research Policy 47, 1554–1567.

Scordato, L., Klitkou, A., Tartiu, V.E., Coenen, L., 2018. Policy mixes for the sustainability transition of the pulp and paper industry in Sweden. Journal of Cleaner Production 183, 1216–1227.

Smith, A., 2003. Transforming technological regimes for sustainable development: a role for alternative technology niches? Science and Public Policy 30, 127-135.

Smith, A., Raven, R., 2012. What is protective space? Reconsidering niches in transitions to sustainability. Research Policy 41, 1025-1036.

Smith, A., Voß, J.-P., Grin, J., 2010. Innovation studies and sustainability transitions: The allure of the multi-level perspective and its challenges. Research Policy 39, 435-448.

Smits, R., Kuhlmann, S., 2004. The rise of systemic instruments in innovation policy. International Journal of Foresight and Innovation Policy 1, 4-32.

Stame, N., 2004. Theory-based evaluation and types of complexity. Evaluation 10, 58-76.

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., Befani, B., 2012. Broadening the range of designs and methods for impact evaluations (downloaded on 30 January 2021 from https://www.oecd.org/derec/50399683.pdf).

Steward, F., 2012. Transformative innovation policy to meet the challenge of climate change: Sociotechnical networks aligned with consumption and end-use as new transition arenas for a low-carbon society or green economy. Technology Analysis and Strategic Management 24, 331-343.

STIPP, 2019. What we do (accessed on 24 February 2021 from https://www.stippsweden.com/about).

Stirling, A., 2009. Direction, distribution and diversity! Pluralising progress in innovation, sustainability and development. STEPS Sussex.

Stufflebeam, D., 2001. Evaluation models. New directions for evaluation 2001, 7-98.

Suurs, R., Hekkert, M., 2012. Motors of Sustainable Innovation: Understanding Transitions from a Technological Innovation System's Perspective, Governing the Energy Transition. Routledge, pp. 152–179.

Suurs, R.A., Hekkert, M.P., 2009. Cumulative causation in the formation of a technological innovation system: The case of biofuels in the Netherlands. Technological Forecasting and Social Change 76, 1003–1020.

Taylor, D., Balloch, S., 2005. The politics of evaluation: An overview, in: Taylor, D., Balloch, S. (Eds.), The politics of evaluation: Participation and policy implementation. The Policy Press, Bristol, UK, pp. 1–20.

Technopolis, 2020a. Meta-evaluation of the second round of strategic innovation programmes after six years (downloaded on 26 June 2022 from https://www.technopolis-group.com/wp-content/uploads/2020/12/metautvardering-av-andra-omgangen-strategiska-innovationsprogram-efter-sex-ar.pdf).

Technopolis, 2020b. Six-year evaluation of the strategic innovation programme BioInnovation (downloaded on 21 March 2023 from https://www.technopolis-group.com/report/six-year-evaluation-of-the-strategic-innovation-programme-bioinnovation/).

Trein, P., Biesbroek, R., Bolognesi, T., Cejudo, G.M., Duffy, R., Hustedt, T., Meyer, I., 2021. Policy coordination and integration: A research agenda. Public administration review 81, 973–977.

Turnheim, B., Berkhout, F., Geels, F., Hof, A., McMeekin, A., Nykvist, B., van Vuuren, D., 2015. Evaluating sustainability transitions pathways: Bridging analytical approaches to address governance challenges. Global Environmental Change 35, 239–253.

Tödtling, F., Trippl, M., 2005. One size fits all?: Towards a differentiated regional innovation policy approach. Research Policy 34, 1203–1219.

Ulmanen, J., Bergek, A., Hellsmark, H., 2022. Lost in translation: Challenges in creating new transformative innovation policy practices. PLOS Sustainability and Transformation 1, e0000031.

van Mierlo, B., Leeuwis, C., Smits, R., Woolthuis, R.K., 2010. Learning towards system innovation: Evaluating a systemic instrument. Technological Forecasting and Social Change 77, 318–334.

Verhees, B., Raven, R., Veraart, F., Smith, A., Kern, F., 2013. The development of solar PV in The Netherlands: A case of survival in unfriendly contexts. Renewable and Sustainable Energy Reviews 19, 275-289.

Vinnova, 2016. Vinnväxt - A programme renewing and moving Sweden ahead (downloaded on 26 February 2021 from https://www.vinnova.se/contentassets/617c28cc186a4f73b9723f16b8820686/vi_16_08t.pdf).

Vinnova, 2018. Cooperation for sustainable innovation (accessed on 24 February 2021 from https://www.vinnova.se/en/m/strategic-innovation-programmes/).

Vinnova, 2019. This is Challenge-Driven Innovation (accessed on 24 February 2021 from https://www.vinnova.se/en/m/challenge-driven-innovation/this-is-cdi/).

Vinnova, 2021a. Projektdatabas (Project database) (accessed on 14 April 2021 from https://www.vinnova.se/sok-finansiering/projekt/?aoName=&current=status&numberofhits=&q=).

Vinnova, 2021b. Strenghening innovation (accessed on 24 February 2021 from https://www.vinnova.se/en/our-activities/).

Vinnova, 2023. Projektdatabas (Project database) (accessed on 14 April 2021 from https://www.vinnova.se/sok-finansiering/projekt/?aoName=&current=status&numberofhits=&q).

Wanzenböck, I., Wesseling, J.H., Frenken, K., Hekkert, M.P., Weber, K.M., 2020. A framework for mission-oriented innovation policy: Alternative pathways through the problem–solution space. Science and Public Policy 47, 474–489.

Weber, K.M., Rohracher, H., 2012. Legitimizing research, technology and innovation policies for transformative change: Combining insights from innovation systems and multi-level perspective in a comprehensive 'failures' framework. Research Policy 41, 1037–1047.

Weible, C.M., Carter, D.P., 2017. Advancing policy process research at its overlap with public management scholarship and nonprofit and voluntary action studies. Policy Studies Journal 45, 22-49.

Weiss, C.H., 1993. Where politics and evaluation research meet. Evaluation practice 14, 93–106.

Weiss, C.H., 1995. Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community-Based Initiatives for Children and Families, New Approaches to Evaluating Community Initiatives: Concepts, Methods and Contexts. Aspen Institute, Washington, pp. 65-92.

Weiss, C.H., 1997. Theory-based evaluation: past, present, and future, in: D. J., R., D., F. (Eds.), Progress and Future Directions in Evaluation: Perspectives on Theory, Practice and Methods. Jossey-Bass, San Francisco, pp. 41-55.

Weiss, C.H., 1998. Evaluation: Methods for Studying Programs and Policies. Prentice Hall, New Jersey, US.

Wesseling, J.H., Meijerhof, N., 2021. Development and application of a Mission-oriented Innovation Systems (MIS) approach. Under review by a scientific journal.

Whitmore, E., 1999. Understanding and Practicing Participatory Evaluation: New Directions for Evaluation. Jossey-Bass, San Francisco, US.

Wieczorek, A.J., Hekkert, M.P., 2012. Systemic instruments for systemic innovation problems: A framework for policy makers and innovation scholars. Science and Public Policy 39, 74-87.

Wise, E., Arnold, E., 2022. Evaluating Transformation – What Can We Learn from the Literature?, Papers in Innovation Studies no. 2022/10. Lund University.

Wise, E., Eklund, M., Smith, M., Wilson, J., 2022. A participatory approach to tracking system transformation in clusters and innovation ecosystems: Evolving practice in Sweden's Vinnväxt programme. Research Evaluation 31, 271–287.

Woolthuis, R.K., Lankhuizen, M., Gilsing, V., 2005. A system failure framework for innovation policy design. Technovation 25, 609-619.

Yap, X.-S., Truffer, B., 2019. Shaping selection environments for industrial catch-up and sustainability transitions: A systemic perspective on endogenizing windows of opportunity. Research Policy 48, 1030-1047.

Yin, R.K., 2009. Case Study Research: Design and Methods. Sage Publications, Thousand Oaks, US.

Zolfagharian, M., Walrave, B., Raven, R., Romme, A.G.L., 2019. Studying transitions: Past, present, and future. Research Policy 48, 103788.

# Appendix A

**Interview guide for BioInnovation SIP**

## Background and context

My PhD research at Chalmers University of Technology is part of the Swedish Transformative Innovation Policy Platform (STIPP), a research project funded by Vinnova that started in 2018. The platform's overall aim is to advance "the understanding of the dynamics and governance of sustainability transitions" (STIPP, 2019) in two transition areas: the transformation to a bio-based economy and the development of smart cities. The platform involves the participation of senior and junior researchers from four Swedish universities: Lund University, Jönköping International Business School, Linköping University, and Chalmers. Specifically, my PhD research focuses on policy evaluation for transformative innovation and change. The rationale behind the project is that current approaches to evaluating innovation policy are not sufficiently adapted to assess transformative change. Thus, its objective is to study how transformative change can be assessed and evaluated (e.g., in relation to specific policy initiatives). Empirically, the research project focuses on the study of the BioInnovation Strategic Innovation Programs (SIP). While this is the target case, I am interested in how the program was developed and what were the initial intended results and outcomes. My point of departure has been so far documents describing the BioInnovation program and its effect logic (see Figure A 1).[39]

## Purpose of the interview

The purpose of the interview is twofold: (i) to understand the main rationales behind BioInnovation and what kind of changes the program was targeting, and (ii) to reflect on the program's impact on sustainability transitions. Note that the purpose of the interview is **not** to evaluate Vinnova as a funding agency or your performance, but rather to reflect on how the program is facilitating changes toward sustainability transitions and how these changes can be evaluated.

[39] Please note that this model of the effect logic corresponds to an old version. An updated version was used in Paper V. The activities and effects listed in the new version used in Paper V, however, just slightly differ to those listed in Figure A 1.

## Format of the interview

The interview will be in English and is expected to last between **one hour and one hour and a half**. It will take place in-person or via Zoom. The interview will be conducted by me, Carolina Haddad. If you agree, I would like to record the interview. The recording will be used for research purposes only and no one outside of our small research team from Chalmers will have access to it. In addition, no other use of the data will be made without your written consent.

## Agenda of the interview

- Introduction—goals of the interview and short introductions
- Setting the scene—background and context of the research
- Questionnaire
    - Understanding the main rationales behind BioInnovation
    - Reflecting on the program's impact on sustainability transitions

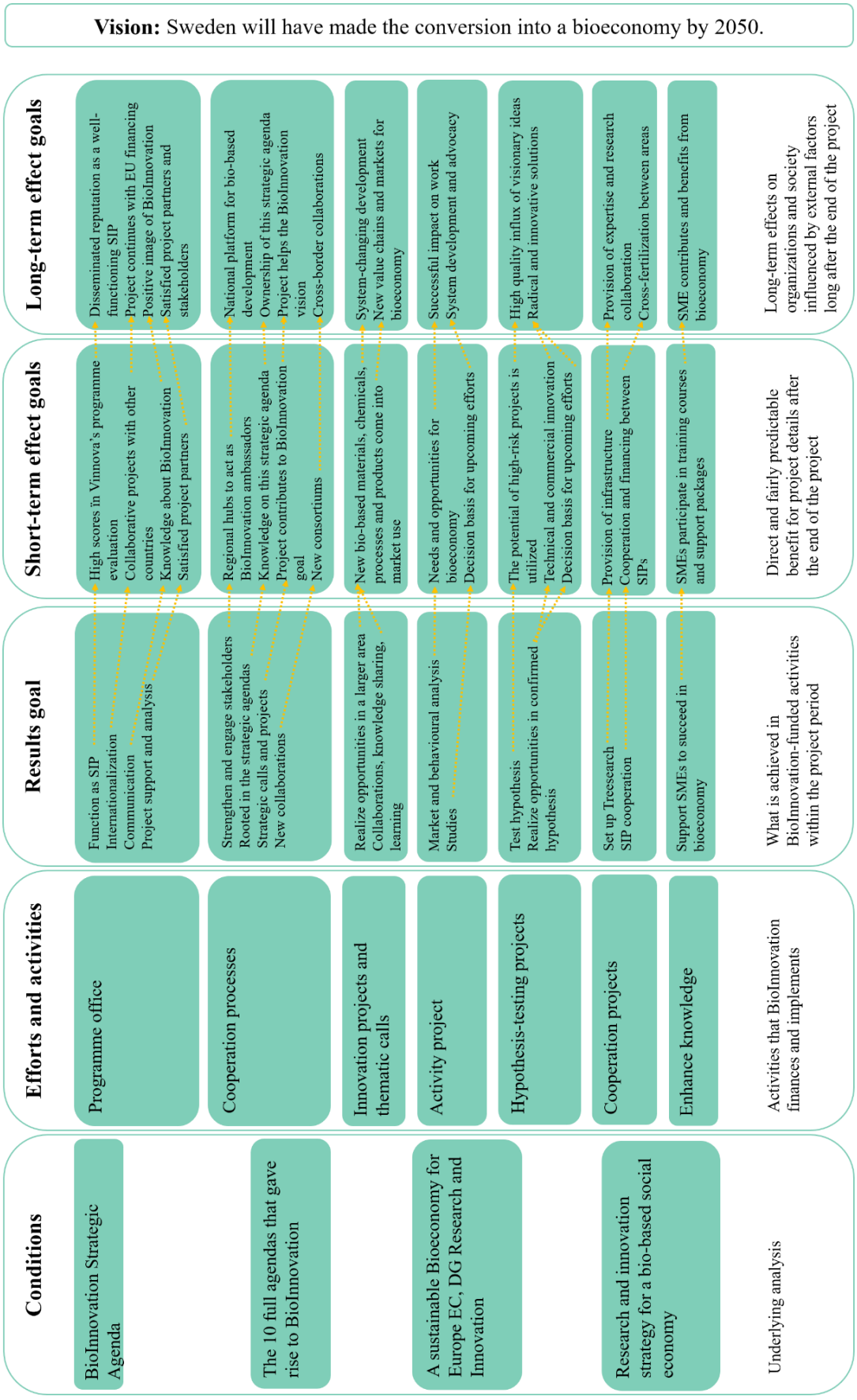


Figure A 1. BioInnovation effect logic. Source: adapted BioInnovation (2014).

**Understanding the main rationales behind BioInnovation**

The BioInnovation SIP emerged under the inspiration of current grand challenges, such as climate change due to fossil emissions, population growth, and the increased consumption of natural resources (BioInnovation, 2020a). BioInnovation reflects on the role of a circular bioeconomy as the way forward for achieving sustainable development and contributing to the Sustainable Development Goals (SDGs). The program works in three priority areas: Chemicals & Energy, Materials, and Construction & Design.

Questions:

1. What is/was your role within BioInnovation?
2. Can you reflect upon the main rationales behind the development of the program? What kind of change did you want to achieve?
3. How was your work with the program's effect logic?
    a. Do you find it useful? Why?
    b. What were the main challenges you faced when developing the effect logic?
4. In relation to the effect logic, (i) what do you mean by the following effect goals? (ii) What were the main activities supporting them (see Figure A 1)?
    a. Cooperation processes and enhancing knowledge?
    b. Cross-border collaborations?
    c. New collaborations vs new consortiums?
    d. System development and successful impact work?
    e. New value chains and markets for the bioeconomy?

**Reflecting on the program's impact on sustainability transitions**

BioInnovation aims to create new business models by encouraging new entrants, as well as by reorienting established actors toward the use of materials with a low environmental impact (BioInnovation, 2020b). Moreover, at the technological level, it envisages the replacement of "climate-impacting materials and chemicals with bio-based alternatives, so-called substitution" (BioInnovation, 2020a, p. 6). There is also the aspiration to use "new bio-based processes and products in integration with existing production infrastructure, and existing processes" to benefit from investment and operations costs (BioInnovation, 2020a, p. 17).

Questions:

5. The following questions relate to key dimensions that support sustainability transitions[40] and aim to reflect what the program expected to achieve (**intended effects**) and what was achieved and how (**outcomes**):
    a. What kind of **knowledge** did the program intend to develop (e.g., market, technologies) and how? What kind of events were there to support the diffusion of knowledge (e.g., workshops, conferences)?
        - Regarding the outcomes, what kind of activities worked and how and what was the role of different actors in realizing them?
    b. How did the program intend to stimulate **experimentation**? What sorts of **uncertainties** did the program target (e.g., technological, market)?
        - Regarding the outcomes, what kind of activities worked and how and what was the role of different actors in realizing them?
    c. How did the program intend to develop new **markets**? Was there any specific market you wanted to develop?
        - Regarding the outcomes, what kind of activities worked and how and what was the role of different actors in realizing them?
    d. How was the **allocation of resources** between different projects? Was it intended to focus on specific technologies?
        - Regarding the outcomes, what kinds of activities worked and how and what was the role of different actors in realizing them?
    e. What types of **infrastructure** did the program intend to develop?
        - Regarding the outcomes, what kind of activities worked and how and what was the role of different actors in realizing them?
    f. How were the **investments** within the different priority areas targeted?
        - Regarding the outcomes, what kind of activities worked and how and what was the role of different actors in realizing them?
    g. Were there any activities targeting the **legitimation** of technologies and/or the program itself in relation to a broader system?
        - Regarding the outcomes, what kind of activities worked and how and what was the role of different actors in realizing them?

[40] These dimensions are part of a general framework that Anna Bergek and I developed to capture key sustainability transition processes (Haddad and Bergek, 2023). This interview will serve as input to test the framework.

h. Was the program expecting to produce any kind of **secondary effects**, for example, in relation to knowledge spill-overs, emergence of intermediary goods and services, or the emergence of pooled labor markets?
    - Regarding the outcomes, what kinds of activities worked and how and what was the role of different actors in realizing them?

i. What did you think about the involvement of different **actors and networks**, for example, big companies within the different sectors, civil society, and SMEs? What were their roles?
    - Regarding the outcomes, were there any specific activities that led to actor and network changes?

j. Did the program target changes in **regulations, rules, or standards**? How was that approached?
    - Regarding the outcomes, were there any specific activities that led to changes in regulations, rules, or standards? What was the role of different actors in realizing them?

6. Summary: Overall, what kinds of changes were you targeting and what has been accomplished in relation to technologies, actors and networks, and regulations, rules, or standards?

| Overview | Intended | Accomplished |
|---|---|---|
| Technologies<br>• New technologies have little chance to break through<br>• Established and new technologies co-exist<br>• Competing new technologies replace established<br>• Established & emerging technologies are combined | | |
| Actors and networks<br>• No change<br>• Established actors re-orient themselves<br>• New actors enter after established actors exit<br>• New entrants outcompete established actors<br>• Established actors & new entrants form alliances | | |
| Regulations, rules, or standards | | |

| | | |
|---|---|---|
| • No change<br>• Limited change<br>• Creation of new regulations<br>• Shocks/crises lead to changes in regulations | | |

7. Can you reflect on the program's main impacts and/or failures? What do you think were the key mechanisms behind the program's success and/or failures? Is there anything that worked for one project but not for others?
8. How do you think BioInnovation contributed to the overall SIP impact goal (see Figure A 2)?
9. Were there any unintended effects due to the program itself (positive or negative)? If yes, can you reflect on which ones and why you think it happened?

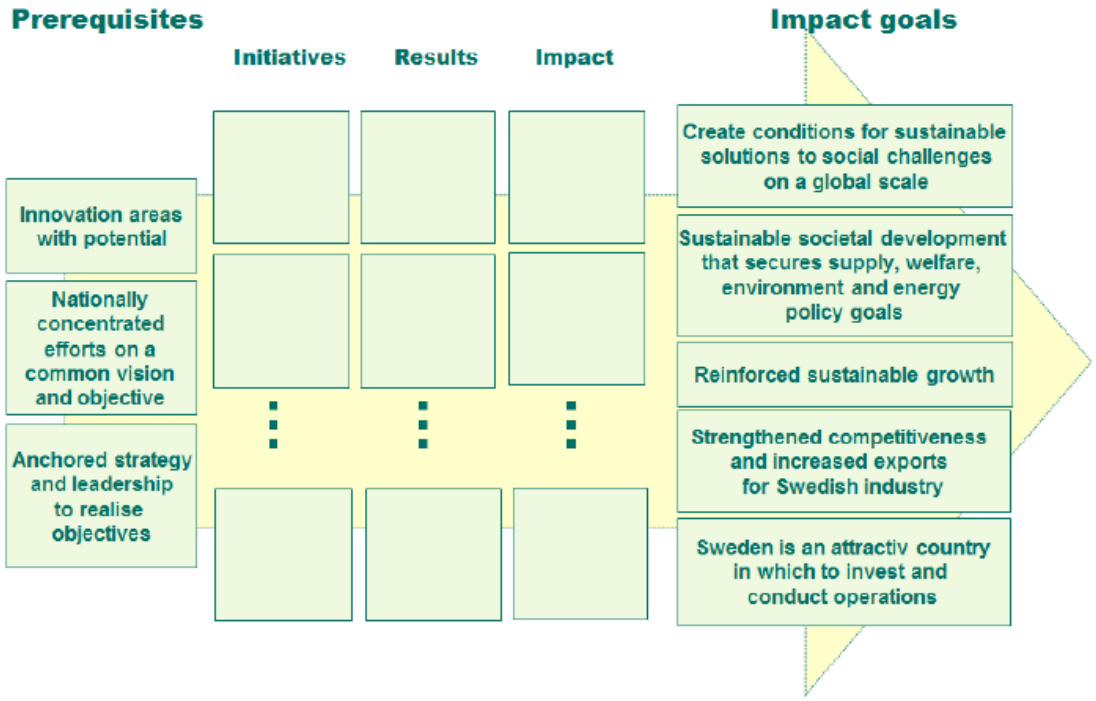


Figure A 2. Strategic innovation program's overall impact goals. Source: adapted from Energimyndigheten (2016).