When theory meets practice in transformative innovation policy evaluation: experiences from Sweden

Downloaded from: https://research.chalmers.se, 2025-01-22 15:36 UTC

Citation for the original published paper (version of record):
Resende Haddad, C., Wise, E., Arnold, E. (2024). When theory meets practice in transformative innovation policy evaluation: experiences from
Sweden. Research Evaluation, 33. http://dx.doi.org/10.1093/reseval/rvae059

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

Research Evaluation, 2024, **33**, rvae059
https://doi.org/10.1093/reseval/rvae059
Article

When theory meets practice in transformative innovation policy evaluation: experiences from Sweden

Carolina R. Haddad [1], Emily Wise [2,*] Erik Arnold [3,4,5]

[1]Department of Technology Management and Economics, Chalmers University of Technology, Gothenburg 412 96, Sweden
[2]CIRCLE and LU Collaboration, Lund University, Lund 221 00, Sweden
[3]Technopolis Group, Brighton BN1 1EE, United Kingdom
[4]Manchester Institute of Innovation Research, Manchester M15 6PB, United Kingdom
[5]Department of Economic History, Lund University, Lund 221 00, Sweden
*Corresponding author. CIRCLE and LU Collaboration, Lund University, Lund 221 00, Sweden. E-mail: emily.wise@fsi.lu.se

Abstract

Government policies increasingly focus on transformative innovation policy (TIP), as programmes are being reoriented towards addressing societal challenges and contributing to sustainability transitions. Evaluation practices need modification to keep up with this change. While a small number of frameworks for TIP evaluation have been proposed, building upon different sustainability transition approaches, little is known regarding the extent to which TIP thinking has been integrated into policy evaluation practice. Hence, there is a need, first, to understand the implications of TIP for evaluation, based on the TIP literature and, second, to investigate the extent to which TIP thinking is used in innovation policy evaluation practice, and the main challenges policymakers face in making evaluation more transformation oriented. To do this, we adapt Edler et al.'s (2012, The Practice of Evaluation in Innovation Policy in Europe', Research Evaluation, 21: 167–82) innovation policy evaluation dimensions to create a framework within which to develop an archetypal TIP programme evaluation, based on the literature. We then investigate the take-up of TIP principles in evaluation in the specific case of Sweden. We use a cross-case comparison of three Swedish innovation programmes, namely the Vinnväxt programme for sustainable regional growth, Challenge Driven Innovation, and the Strategic Innovation Programmes, to analyse the extent to which innovation policy evaluation practice aligns with this archetypal framework. Finally, we identify three challenges policymakers face when trying to reorient evaluations towards being more aligned with this archetypal framework.

Keywords: transformative innovation policy; evaluation; programme evaluation; policy programme.

1. Introduction

The emergence of transformative innovation policy (TIP) has triggered the development of new programmes and initiatives targeting societal challenges (Borrás and Schwaag Serger 2022; Schwaag Serger and Palmberg 2022). Examples include the Swedish Strategic Innovation Programmes (SIPs), the Dutch Top Sector strategy and Challenge Finland (Janssen 2019; Schwaag Serger and Palmberg 2022). These initiatives are designed to use aspects of TIP to enable system innovation, as framed in the sustainability transitions literature (Geels 2004; Schot and Steinmueller 2018). Such aspects include the focus on addressing Grand Challenges, the need to develop purposeful and directional innovation, and the facilitation of system-wide transformation, bringing about changes at multiple structural and systemic levels (Diercks, Larsen and Steward 2019; Borrás and Schwaag Serger 2022). Additionally, TIP promotes intervention at all levels of governance and across policy mixes, and focuses on the involvement of multiple actors (Haddad et al. 2022).

There is now an increasing need to adapt evaluation practices for TIP (Arnold et al. 2018; Haddad et al. 2022; Janssen, Bergek and Wesseling 2022) to reflect shifts in innovation policy thinking (Molas-Gallart and Davies 2006; Haddad and Bergek 2023). To address this, several frameworks have been introduced in the literature. Some build on transition frameworks, such as the multi-level perspective (MLP) and the technological innovation system (TIS) functions, to analyse socio-technical system change ex-ante (Kern 2012; Bos, Hofman and Kuhlmann 2016), to suggest additional functions for 'creative destruction' in policy mixes for transitions (Kivimaa and Kern 2016), or to evaluate the contributions of policies to building up TIS (Janssen 2019). Other authors propose principles (Molas-Gallart et al. 2021) or step-wise frameworks (Haddad and Bergek 2023) for evaluating TIP interventions. Yet, others suggest schemes for evaluating the effects of transition experiments (Luederitz et al. 2017; Williams and Robinson 2020).

However, little is known regarding the extent to which TIP thinking has been integrated into evaluation practice. Some authors argue it gets lost in translation (Diercks 2018; Ulmanen, Bergek and Hellsmark 2022; Bergek, Hellsmark and Karltorp 2023) so practice lags behind theory, especially in evaluating innovation policy (Molas-Gallart and Davies 2006; Magro and Wilson 2013).

Edler et al. (2012) analysed innovation policy evaluation in 25 European countries, and found that, although practice includes both formative and summative elements, it is mostly summative evaluation that influences higher-level policy changes, while formative evaluation triggers only minor modifications. Most evaluations focus on economic and technological impacts; few include the effects of demand-oriented policies and innovation diffusion measures. More recently, Borrás and Laatsit (2019) investigated the extent to which the 28 EU countries did system-level policy evaluation. They concluded that only a few countries have developed system-level evaluations (including The Netherlands, Sweden,

© The Author(s) 2024. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

Austria, Finland, Ireland, and Germany). This limits many countries' learning about their innovation systems and policy dynamics.

To the best of our knowledge, no published study has yet explored how TIP characteristics are translated to a programme evaluation strategy[1] in practice. This paper aims to bridge this gap by addressing three research questions:

i) What are some of the key characteristics of a TIP programme evaluation strategy according to the literature?
ii) To what extent does current innovation programme evaluation practice match theoretically derived characteristics of a TIP evaluation strategy?
iii) What are the main challenges involved in adapting current innovation programme evaluation practice to TIP?

To address these questions, we start with the Edler et al. (2012) framework for describing evaluation practice in second generation innovation policy (Arnold et al. 2018) and adapt it to third generation (TIP) characteristics (based on the literature) to build an evaluation strategy for TIP. We then use this TIP evaluation strategy framework to do a cross-case comparison of three Swedish innovation programmes to analyse the extent to which current innovation policy evaluation practice aligns with what is proposed in the TIP literature. Although these programmes—Vinnväxt, the Challenge Driven Innovation (CDI) programme, and the Strategic Innovation Programmes (SIPs)—were not originally designed to be transformative, they have been re-oriented over time to address societal challenges and have previously been analysed from a TIP perspective (Grillitsch et al. 2019; Borrás and Schwaag Serger 2022; Wise et al. 2022; Åström and Arnold 2023). These programmes are long-term and involve multiple stakeholders, fulfilling some TIP characteristics (Haddad et al. 2022). Therefore, they serve as interesting starting points to analyse current innovation policy evaluation practice, given that Sweden is one of the few countries that have developed system-oriented evaluation practices (Borrás and Laatsit 2019).

The paper is structured as follows. In Section 2, we adapt and amend the evaluation dimensions of Edler et al. (2012) into an analytical framework and reflect on how they differ between traditional and transformative innovation policy. Section 3 describes the research design and introduces the cases covered in the paper. Section 4 presents the evaluation strategies of the three cases and the main findings of a cross-case analysis. Sections 5 and 6 discuss and conclude, respectively.

2. Analytical framework

To analyse the evaluation practices of the three Swedish innovation programmes empirically, we need a framework that covers the main dimensions of an evaluation strategy for TIP. We, therefore, rely on the four evaluation dimensions proposed by Edler et al. (2012): (1) evaluation set-up, which covers the timing of the evaluation, its purpose and who commissioned it; (2) policy measures, which relate to the policy objective and target groups (which we merged with the first dimension); (3) main questions, which is divided between the topic covered in evaluation (e.g. input/output/outcome) and the type of impact these topics (such as socio-technical, environmental, or economic) are expected to cause; and (4) methods used in the evaluation in terms of data collection and data analysis.[2]

We use this framework as it comprises the 'first steps towards a typology' of the key characteristics of innovation policy evaluations (Edler et al. 2012: 168) and is also the basis of the current Science and Innovation Policy Evaluation Repository (SIPER 2023). To the best of our knowledge, Edler et al.'s framework is one of the most comprehensive approaches to reflect innovation policy evaluation characteristics. Therefore, it serves as a reasonable starting point to adapt the characteristics of what would characterize as a second framing of innovation policy evaluation to a TIP evaluation strategy. Accordingly, we tweak some of its dimensions to reflect the insights from the TIP literature. Additionally, Edler et al.'s framework covers most of the characteristics of evaluation, as compared for example with previous attempts to categorize evaluations found in the general evaluation literature (see e.g. Stufflebeam 2001; Alkin 2013). These attempts, for example, include categories such as timing, scope, purpose, and methods, all of which are also covered by Edler et al., who also tailor such categories to innovation policy.

Table 1 summarizes the key evaluation dimensions and categories and their connection to TIP. We treat these dimensions and categories as an archetypal TIP evaluation strategy and consider the extent to which individual programmes align with them, as opposed to what the literature would consider more traditional and non-systemic approaches to STI evaluation strategy. In the subsections below, we elaborate on how each category differs from TIP in relation to these previous approaches.

2.1 Evaluation set-up

Regarding evaluation set-up, we removed the commissioner category from Edler et al. (2012), because this does not change in the context of TIP, and added the scope of evaluation and stakeholder involvement, to highlight the importance of involving a broader set of stakeholders in evaluation and learning, beyond traditional R&I actors (cf Haddad et al. 2022).

The evaluation set-up dimension comprises five main categories: timing, purpose, scope, level of analysis, and stakeholder involvement. Regarding *timing*, evaluation can be conducted *ex-ante*, i.e. before the implementation of an intervention, or *ex-post*, after the implementation of the intervention (Khandker, Koolwal and Samad 2010). It can also occur during the implementation, in the form of monitoring and interim evaluations. Some also propose a 'real-time' approach for evaluating R&I projects and programmes, in which the vision of the future and pathways for change are constantly adapted and reconsidered to reflect intermediate results (Matt et al. 2023).[3] More traditional and non-systemic approaches to research and innovation funding would develop evaluations ex-ante and (mainly) ex-post (Arnold 2004; Magro and Wilson 2013) since such interventions are based on static theories of change. For TIP, there is a need for an evaluation strategy that covers processes of monitoring, interim evaluation and learning since the interventions and their contexts are dynamic and complex (Wise and Arnold 2022). They could also benefit from a 'real-time' approach.

The *purpose* of evaluations can be summative, assessing policy for accountability, or formative, aiming to improve policy by supporting learning (Knill and Tosun 2012).

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

Table 1. Key evaluation categories and their connection to transformative innovation policy.

Dimensions[a]	Categories	Description	Traditional approaches		TIP
Evaluation set-up	Timing	When in the programme's life cycle evaluation occurs, i.e. before, during and/or after the programme has been implemented.	Mainly ex-post	⟷	Holistic evaluation strategy including stronger ex-ante and real-time monitoring components
	Purpose	Why the evaluation is developed and by whom it is used.	Summative (accountability)	⟷	More formative and developmental
	Scope of the evaluation	Clarifies what will be evaluated (in terms of instruments and programmes) and the targeted system it aims to change.	Single policy instruments	⟷	Elements of the policy mix (across instrument mix and policy areas) targeting changes on broader socio-technical systems
	Level of analysis	The level at which the evaluator will analyse the effects and impact of an intervention, e.g. at the level of projects, portfolio of projects and/or programmes.	Intervention at one level	⟷	Nested approach across multiple levels
	Stakeholder involvement	The actors involved in the evaluation, e.g. intervention beneficiaries, funders, experts, civil society, etc.	Within-government agents (e.g. project participants and/or programme managers)	⟷	Involvement of a broader range of stakeholders
Main questions	Input/output/outcome	What is being evaluated: inputs (e.g. human and financial resources), outputs (first level and concrete results of a policy intervention, e.g. number of patents and publications), and/or outcomes (the medium-term consequences of an intervention that relate to the programme goals or aim).	Input-output indicators	⟷	Transformative processes
	Additionality	The extent to which something additional happened due to policy intervention, that would not have occurred without it. (cp 'net effect' or 'counterfactual').	Input-output additionality, firm-level behavioural additionality	⟷	System-level additionality
	Directionality	The extent to which technical change addresses specific problem areas and solves specific societal problems	Not very important, on the assumption that all technical change is potentially beneficial	⟷	Very important, since the purpose of policy is to address specific societal problems
Methodology	Causality (how to address causality)	The relationship between cause and effect and the different approaches to establish it.	Successionist causality	⟷	Complex causality

[a] Dimensions and categories adapted from Edler et al. (2012) to reflect TIP characteristics. Arrows (⟷) represent a spectrum between more traditional and non-systemic evaluation approaches and what would comprise a TIP evaluation strategy according to the literature.

Evaluations can also be developmental, focusing on 'adaptive development', successively improving the intervention based on changes in context, the client, or the emergence of alternatives (Patton 2010). Traditionally, STI policy evaluations have been summative but formative evaluations are more important in system-oriented innovation policy evaluation (Borrás and Laatsit 2019). In TIP, the emphasis has so far been on formative and developmental evaluation (Wise and Arnold 2022).

In traditional STI policy, the *scope of the evaluation* is typically of a single policy instrument or programme (Arnold 2004; Flanagan, Uyarra and Laranja 2011), and not the system it aims to change. As shown by Borrás and Laatsit (2019) for the European case, only a few countries go beyond the analysis of specific instruments to also consider the system it aims to change. When the system is considered, they are often targeting the evaluation of innovation system performance (e.g. of specific regional, technological or national

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

innovation systems). In contrast, the TIP literature tries to account for complex policy mixes (across different policy areas[4]), interactions among instruments (sometimes bundled in programmes), synergies, and conflicts (Janssen 2019; Haddad et al. 2022). In TIP, these instruments target changes in socio-technical systems, implying broader systems boundaries (Haddad and Bergek 2023). Therefore, considering the type of system and the systems' boundaries becomes important when defining the scope of TIP evaluations.

Concerning *level of analysis*, the spectrum differs between analysing the effects and impact of individual projects or experiments, at one level and timing, or having a nested and cumulative approach to the evaluation. Traditionally, research and innovation evaluation is done at the micro-level, i.e. individual projects or collections of projects (Arnold 2004; Molas-Gallart and Davies 2006). TIPs, using nested and cumulative approaches, can span multiple levels, from specific niche experiments and initiatives and projects to programmes and the combination of programmes over time (Molas-Gallart et al. 2021).

Stakeholder involvement in traditional STI evaluation would involve actors close to the programme (Magro and Wilson 2019), such as project participants, i.e. the direct beneficiaries, and/or programme managers. TIP calls for a broader involvement of stakeholders, including, among others, policymakers, service providers, academics, and civil society organizations (Arnold et al. 2018; Magro and Wilson 2019). This means that a broader range of stakeholders will be affected and/or affect the intervention itself and, hence, they should also be considered in evaluations. Consequently, evaluations should be more inclusive and move more stakeholders to the centre of the evaluation process, where they also evaluate themselves (Amanatidou et al. 2014; Magro and Wilson 2019; Haddad et al. 2022). The involvement of different stakeholders will also depend on the context in which the evaluation is developed. This is because the engagement of additional actors in the evaluation can also lead to conflicts and demand capacity that is not always available from those performing the evaluation. Stakeholder involvement might be more difficult in countries where programmes are designed top-down and do not involve citizens or other non-obvious actors in the policymaking process.

2.2 Main questions

For the main question dimension, which covers both topic and impact dimensions from Edler et al.'s (2012) original framework, we focus on evaluation categories that differ between traditional and TIP evaluation, including input/output/outcome, additionality categories, and directionality (Amanatidou et al. 2014; Janssen 2019; Haddad and Bergek 2023).

We sub-divide the *input/output/outcome* category between a focus on input-output measures versus transformative processes (or outcomes). Traditional innovation policy evaluations focus on input-output measures such as money spent in collaboration (input) and the number of outputs resulting from the collaboration (output) (Amanatidou et al. 2014). In the TIP literature, 'transformative processes', in contrast, refer to innovation and transition processes expected to enable transitions and for which TIPs should strive (Luederitz et al. 2017; Molas-Gallart et al. 2021; Haddad and Bergek 2023). Some authors have proposed frameworks for identifying socio-technical system outcomes, for example, by covering processes at the niche, regime (and landscape) levels (Kern 2012; Ghosh et al. 2021), or by broadening the functions of the innovation system (cf Kivimaa and Kern 2016; Bergek and Haddad 2022; Elzinga et al. 2023).[5] This is also closely related to the type of systemic change the intervention intends. In TIP, this is broadened towards addressing change in socio-technical systems that fulfil a societal function (Bergek and Haddad 2022), as defined in the sustainability transitions literature (Geels 2004).[6]

Regarding transformative processes, Bergek and Haddad (2022) adapt the TIS functions to cover processes related to sociotechnical system change, as well as structural changes in actor networks and institutions.[7] Elzinga et al. (2023) use the functions as a basis to outline processes that can be used to assess missions, encompassing programming functions such as providing directionality, and functions related to inducing innovation and destabilizing the regime, e.g. related to resource mobilization/withdrawn and market formation/market destabilization. Ghosh et al. (2021) list a range of 12 transformative outcomes, related to building and nurturing niches, expanding, and mainstreaming niches, and opening up and unlocking mechanisms.

Innovation policy has traditionally relied on calculations of input and output *additionality*, respectively the additional effect of public funding on total firm R&D investments, and the additional output (e.g. in terms of patents, publications, and sales) achieved by public support (Georghiou and Clarysse 2006; Clarysse et al. 2009). Another type of additionality, namely behavioural additionality, measures the additional effect of public interventions on the behaviour of firms and other actors (Georghiou and Clarysse 2006; Amanatidou et al. 2014). The concept of behavioural additionality has also been expanded to capture effects of policy on learning (Clarysse et al. 2009). Janssen (2019) suggests stretching the concept of additionality to look at how functions are supporting innovation systems creation. In other words, this entails understanding how policy is contributing to building up TIS in terms, for example, of strengthening functions such as entrepreneurial experimentation (e.g. new entrants in the system), or resource mobilization (e.g. increase in R&D funding and R&D staff).

Haddad and Bergek (2023) suggest a further stretching to address system-level additionality in the form of contributions to changes in transformative processes [e.g. those proposed by Ghosh et al. (2021) and Bergek and Haddad (2022)], which consider, for example, both innovation and transition processes, including those related to regime destabilization. This is different from what has been proposed before, as it suggests that additionality is not only about attributing the effects of policies for accountability purposes, but also about supporting policy learning and adaptation. This would mean assessing the extent to which policy programmes are able to support change in key transformative processes. As will be highlighted in the methodology dimension below, analysing the contribution of programmes in promoting changes in entire sociotechnical systems is the additionality that evaluators should be looking for. This will require alternative ways to identify causality and will also be context specific.

In terms of *directionality*, TIP emphasizes the importance of directing change and innovation towards enabling transitions, which was not previously covered in earlier approaches to innovation policy (Weber and Rohracher 2012; Diercks, Larsen and Steward 2019). Therefore, rather than supporting the

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

development of more and faster innovations, policies should be oriented towards addressing societal challenges (Edler and Boon 2018). Setting a direction could involve multiple 'acceptable development paths' (Weber and Rohracher 2012) or multiple directions based on different problem and solution constellations (Wanzenböck et al. 2020). Nonetheless, policymakers should prioritize options that contribute to solving societal challenges over alternative paths that can lead to unsustainable options (Schot and Steinmueller 2018; Bergek, Hellsmark and Karltorp 2023).

2.3 Methodology

In relation to the methodology dimension, we focus on *causality*, given that for TIP the complexity of both the interventions and the systems in which they intervene make it especially difficult to isolate the effect of policy on transformative change (Arnold et al. 2018; Janssen 2019). A crucial role of evaluations is to explain the causal links between interventions and outcomes (Weiss 1995; Pawson and Tilley 1997; Funnell and Rogers 2011), and a theory of change (ToC) should also therefore be considered part of the policy design (Molas-Gallart et al. 2021). Evaluators of transitions and missions are converging on the view that ToC-based approaches are necessary, while specific evaluation tools are not necessarily different between traditional and TIP evaluations (Arnold and Wise, Forthcoming). Theories of change should encompass both the logic model, which specifies the sequence of steps from inputs and activities to outputs and outcomes, and the causal assumptions behind it, detailing why and how the links of the logic model are supposed to work (Mayne 2015). They also need to reflect contextual conditions, as initiatives targeting transformation are influenced by, among other aspects, the programme's staff, its organizational setting, context structures (e.g. sectoral and political context), as well as exogenous forces, such as landscape pressures (Haddad and Bergek 2023).

The traditional approach to impact evaluation is to try to compare the additional effects of a policy intervention to the counterfactual, i.e. what would have happened in the absence of the intervention (Bartle and Morris 2010; White 2010). Attempts to quantify the additionality of R&I interventions follow this notion, aiming to calculate the 'bang for the buck' of R&D measures (Janssen 2019). However, these approaches are rarely useful outside simple interventions, say little about how and why change happens (Pawson and Tilley 1997), and are not equipped to explain long-term transformative change (Geels 2022).

When it comes to complex interventions (such as TIP), there is a need for alternative approaches for establishing causality (Hind 2010). Given the need to incorporate the analysis of transformative processes and account for system-level additionality, mechanism-based approaches, e.g. process tracing, can serve as an alternative to explain causality in complex programmes (de Oliveira, Lacerda and Negro 2020; Haddad and Bergek 2023). Geels (2022) suggests that transition frameworks, such as TIS, MLP and strategic niche management (SNM), resonate well with critical realist philosophy, which emerged as an alternative to positivism and sees reality as based on three realities: empirical (what is observed and experienced), actual (causal mechanisms generating events), and real (entities and structures with causal power). Such ontology follows a generative view of causality, in which outcomes are explained by underlying mechanisms acting in a specific context (Pawson and Tilley 1997). Geels (2022) also argues that generative causality accommodates multiple types of causation, including complex causality (e.g. conjunctural, configurational, and event-chain causality).[8] This might imply that TIP evaluations should consider using complementary mixed methods approaches 'that can accommodate various forms of complex causality' (George and Bennett 2004: 5).

3. Research design

In this section, we detail the multiple-case design approach (Bryman and Bell 2015) we used to explore the extent to which current evaluation practices are aligned with TIP characteristics, focusing on three long-running programmes funded by Sweden's innovation agency, Vinnova.

3.1 Selection and introduction of cases

Vinnova develops and implements programmes that contribute to its mission of building Sweden's innovation capacity and contributing to sustainable growth. Since its presidency of the European Union in 2009, Sweden has shifted policy attention to addressing societal challenges. Already in 2016, the OECD acknowledged Vinnova's swift launch of the challenge-driven innovation programme and other efforts to shift the focus of research and innovation funding to address societal challenges (OECD 2016). In recent years, Vinnova has been actively engaged in the Transformative Innovation Policy Consortium (TIPC) and OECD working groups that are exploring and experimenting with new approaches to implement transformative and mission-oriented innovation policy agendas. This work is reflected not only in new interventions such as mission-oriented experiments in food and mobility systems (Larrue 2021; Velasco and Witte 2022; Boni et al. 2023) and the ongoing design and implementation of Impact Innovation (Vinnova 2022a), but also in the adaptation of existing interventions to tackle newer policy objectives. We analyse three Swedish innovation programmes: Vinnväxt, CDI and SIPs. Although they were not designed to be transformative from the start, the cases were chosen based on their system-level logic and adoption of transformative aims (particularly in the most recent generations of these interventions). The cases also have characteristics worth analysing from an evaluation perspective. Thus, these cases fulfil an illustrative function in exploring the alignment of current evaluation practice with TIP characteristics. A description of each programme follows.

3.1. 1 Vinnväxt

Vinnväxt (initially launched in 2001) promotes sustainable growth in Swedish regions by engaging triple helix actors in long-term collaborative initiatives to develop internationally competitive regional innovation systems in specific growth areas. With long-term (up to 10 years) funding, the regions must be able to coordinate and develop new patterns for collaboration (between business, academia and the public sector) that lead to more efficient innovation processes and more effective system-level action. Vinnväxt has reoriented itself over time to support the Sustainable Development Goals (SDGs), with most recent initiatives addressing challenges in the areas of digital health innovation, green chemistry and agricultural technology. In addition, Vinnväxt has presented 'evolving monitoring and evaluation practices', focusing on

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

Table 2. Vinnväxt initiatives (by starting points) over time.

Vinnväxt 2003	Vinnväxt 2004	Vinnväxt 2006/8	Vinnväxt 2013	Vinnväxt 2016	Vinnväxt 2019
Robotic Valley	Fiber Optic Valley	Processum	GeoLife Region	Automation Region	AgTech 2030
Skåne Food Innovation Network	New Tools for Health	Biorefinery Cluster	Paper Province	Urban Magma	DigitalWell Arena
Uppsala BIO	Göteborg BIO	Peak Innovation	Smart Housing Småland	Visual Sweden	Climate-smart Process Industry
	Process IT Innovations	Printed Electronics Arena			
	Triple Steelix	Smart Textiles			

Source: Vinnova.

learning and reflexivity and adapted to cover system-level effects (Wise et al. 2022: 276).

Vinnväxt has had six programme calls since 2002—resulting in six 'generations' of Vinnväxt initiatives—see Table 2. To date, 21 Vinnväxt initiatives have been funded, of which 10 have 'graduated', and 3 have not proceeded through all phases.[9] As of 2023, there are 8 ongoing initiatives. Each Vinnväxt initiative is expected to have a collaborative vision and strategy in line with the region's priorities (as set out in smart specialization strategies). The programme provides funding of some €200–800 thousand p.a. and other support services over a period of (up to) 10 years. On average, Vinnväxt initiatives crowd in ~4 times Vinnova's investment (Kontigo 2016), i.e. they attract 4 times as much complementary funding as Vinnova supplies via Vinnväxt, and manage portfolios of innovation and system-development activities with total budgets in the range of €1.3–45 million p.a.

3.1. 2 Challenge-driven innovation

The CDI programme was launched in 2011. CDI funds visionary projects that address challenges formulated by the organizations submitting applications, and that contribute to achieving one or more SDGs. CDI project consortia are expected to encompass skill sets and engagement from the political sphere, industry, the public sector, civil society and research. It is critical that the different actors involved in the efforts demonstrate a user and needs perspective, as well as the ability to disseminate and scale up the solutions/innovations that are developed. With its strong demand-side perspective and challenge-oriented directionality, CDI has been classified as a mission-oriented policy initiative (Larrue 2021).

CDI provides project funding in three stages: (1) initiation, focusing on developing an idea and planning its development (up to 9 months); (2) collaboration, developing the partnership further, starting to develop and test innovative solutions in greater depth (up to 2 years); (3) scaling up the results and building the business model (up to 2 years) (Ramboll 2020). These involve increasing levels of state funding, but also increased co-funding from the consortia over time (for a maximum of 5 years). As of December 2021, CDI had funded 783 projects. Just over 500 had received Stage 1 funding, and just over 50 had received funding for Stage 3. More than 2,500 unique organizations had participated in CDI, and Vinnova had granted a total of almost €250 million in funding.

3.1. 3 Strategic innovation programmes

The 2012 Research Bill triggered Vinnova to invite interested groups of research and innovation actors to propose national innovation agendas, generating wide interest among potential participants in establishing new innovation programmes. The SIPs aim to implement innovation agendas proposed by public-private partnerships in areas of industrial and strategic importance to Sweden, while increasing the country's international competitiveness and developing solutions to address global societal challenges (Vinnova 2022c). They are the most recent effort from Vinnova (together with the Swedish Energy Agency and the Formas Swedish Research Council for Sustainable Development,) in developing innovation programmes targeting system-wide transformation.

Around 136 agendas were eventually produced. From 2013 on, Vinnova, Formas and the Swedish Energy Agency, invited proposals for SIPs, which could build on these agendas or other sources. A total of 17 SIPs were funded (Table 3) for up to 12 years each, depending on positive interim evaluations done every 3 years. Each SIP runs multiple projects and receives €2–8 million per year in state subsidy. Most projects must be co-funded by project participants, which almost doubles the amount invested.

3.1. 4 Governance and evaluation of the programmes

As Fig. 1 illustrates, CDI projects are governed directly by the funder, while in each of the Vinnväxt initiatives and each SIP there is an intermediary organization (often but not always a public-private partnership) that writes the initial proposal, represents a cluster of stakeholders, manages the initiative or SIP and a programme office, and is accountable to the funders. All three programmes provide multi-stage funding, so the chief evaluation focus is naturally on the performance (both in terms of innovation that addresses a challenge but also in terms of systemic action and transformation of the system) of the entities that get the money. Evaluation for policy learning, on the other hand, focuses on the programme level.

3.2 Data and methods

This paper uses a cross-case analysis research design (Eisenhardt 1989), based on multiple sources. We conducted 8 semi-structured interviews in May–June 2022 with policymakers and practitioners involved in the evaluations of each programme (see Table 4). Each interview lasted 80–120 min, aiming to understand details of the programmes' evaluation strategies.

We reviewed: (1) programme descriptions and the initial calls for proposal, announcing competitions to take part in the CDI, Vinnväxt and SIP programmes (2) call texts on the programmes' or funders' websites[10] announcing competitions open to existing programme participants, (3) the programmes' evaluation reports produced by external evaluators, such as Vinnväxt's (after 3 and 6 years) (Vinnova 2021a), SIPs' interim evaluations (after 3, 6 and 9 years) (Vinnova 2022b), and CDI analysis reports (Ramboll 2020), (4) Vinnova's internal documents guiding the process of evaluation, including self-report questionnaires and agenda items for annual reconciliation meetings between programme managers and Vinnova (Vinnväxt and SIPs), and (5) the calls for external programme evaluation via public procurement.

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

Table 3. SIPs (by starting points) over time.

SIPs 2013	SIPs 2014	SIPs 2015	SIP 2017
Metallic Materials	BioInnovation	Drive Sweden	Viable Cities
LIGHTer	Innovair	InfraSweden 2030	
Sustainable Production in Sweden	IoT Sweden	Medtech4Health	
Process industrial IT and automation	Grafen	RE: Source	
	Smarter Electronic Systems	Smart Built Environment	
Swedish Mining Innovation	Swelife		

Source: Vinnova (2021b).

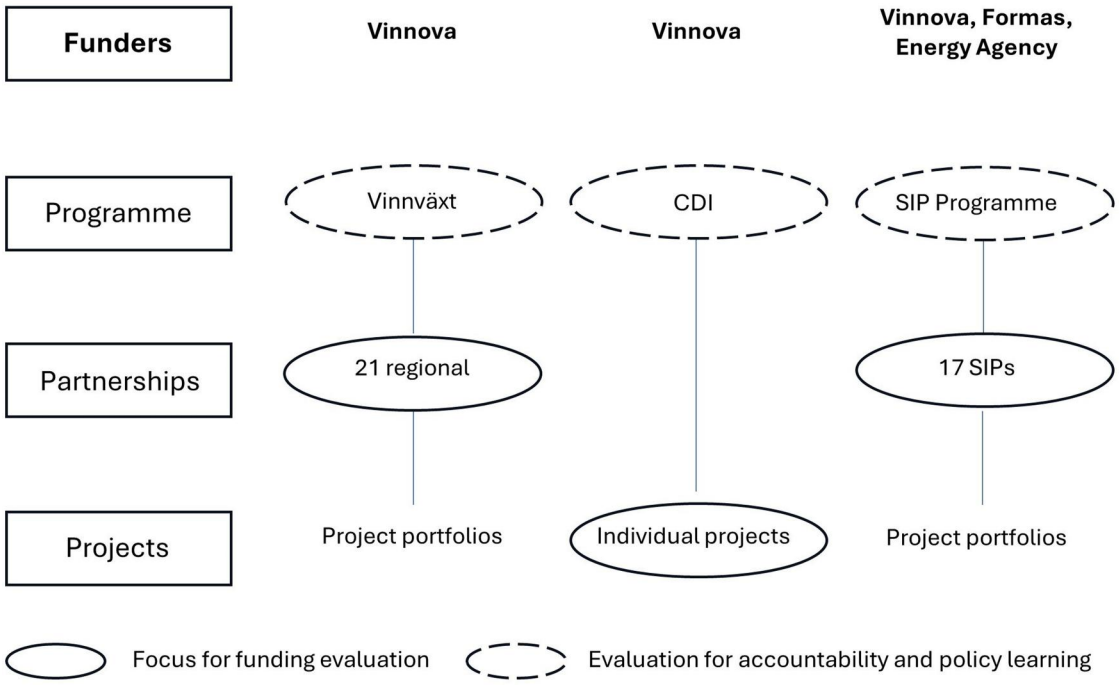

Figure 1. Governance and evaluation levels of the programmes.

Table 4. List of interviews.

No.	Programme	Role	Type of organization
1	Vinnväxt	Policy analyst	Vinnova
2	Vinnväxt	Practitioner involved in evaluation	Academia
3	Vinnväxt	Practitioner involved in evaluation	Academia
4	CDI	2 policy analysts	Vinnova
5	CDI	Practitioner involved in evaluation	Consultancy
6	SIPs	3 policy analysts	Vinnova
7	SIPs	Practitioner involved in evaluation	Consultancy
8	SIPs	Practitioner involved in evaluation	Consultancy

Information gathered in interviews and review of documents were summarized in three extensive case reports (one for each programme) that synthesized programme aims and provided a detailed description of the evaluation strategy (elaborating purpose, uses and approaches taken for ex-ante, monitoring, interim and ex-post evaluation). We then analysed the cases using the framework proposed in Section 2. Specifically, two authors scrutinized and coded manually, via document analysis (Bowen 2009), the case reports to identify whether and how the evaluations addressed each category of the analytical framework (highlighting each category as a comment in the case reports). Earlier drafts of the three cases were then revised by the third author, as well as by the interviewees. Additionally, the cases were discussed in two half-day workshops held in October and November 2022 with Vinnova (respectively, 7 and 6 Vinnova representatives participated), in which the three authors led discussions of the main findings from the cases and of lessons learnt from current evaluation practice in the three programmes. We used our different data sources as well as feedback from interviewees and insights from the workshops for data triangulation.

We then performed a cross-case analysis of the programmes. First, we established by discussion a specific three-point scale—low/non-existent, medium, or high—for the degree of alignment between each evaluation category and the TIP archetypal framework (see Supplementary Appendix A).

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

Then, based on these scales, we discussed and agreed, category by category, how well each programme's current evaluation practice would align with a TIP evaluation strategy.

Finally, we identified key cross-cutting challenges relevant for aligning evaluation practice more closely to an archetypal TIP evaluation strategy. These were identified based on discussions with Vinnova during the two workshops and on our own assessment of key TIP evaluation issues. These challenges also led us to make improvements in the original analytical framework. Specifically, we included the importance of setting the boundaries of the evaluation within the scope of evaluation dimension, given that TIP broadens the scope considerably compared to traditional STI policy evaluation (see Section 2.1).

4. Case analysis: Vinnväxt, CDI and SIPs

A brief description of each programme's evaluation strategy follows, with an overview of main findings presented in Table 5 below. Detailed case descriptions can be found in Supplementary Appendix B. Although the three programmes exhibit TIP-like ambitions and characteristics, none can claim to be 'fully TIP'. As such, we do not expect that their evaluation strategies and existing practices will be fully aligned with an archetypal TIP evaluation strategy (as portrayed in the framework in Section 2). However, we do hope to identify existing practices that would be suitable to build upon, as well as priority areas for new development in evaluation practice for the new generation of transformative innovation programmes.

4.1 Description of evaluation strategies

4.1. 1 Vinnväxt

The Vinnväxt evaluation strategy encompasses a mix of purposes and approaches over the course of each initiative's 10-year implementation period, with a strong focus on learning and development throughout. The evaluation includes the use of programme and initiative level logic models to establish directionality (ex-ante), a robust set of continual monitoring activities (involving stakeholders) to promote continual learning and development over time, regular interim external evaluations (every 3 years) and a final ex-post evaluation (through self-assessment exit reports and an external impact assessment). The programme directionality highlights expectations both for new knowledge and innovation (frame 2), and for more effective actor interactions and renewal of the (regional) system—addressing broader societal challenges (frame 3). As such, the evaluation strategy tracks traditional indicators of innovation and key events/processes of systems change. A key feature of the Vinnväxt evaluation strategy is the palette of continuous monitoring activities including active strategic dialogue between Vinnova and the funded initiatives, and the investment of at least five percent of Vinnova funding for strategic and reflexive learning (where external action researchers or consultants provide external 'constructively critical' coaching and supporting analysis for the initiative's management and for the board/steering group, as well as documenting the story of the initiative and the contributions it has made to transforming the system).

4.1. 2 Challenge-driven innovation

The evaluation and learning practices within CDI have been an embedded part of the stage-gate design—ie that an assessment of results achieved and forward-looking strategy occurs as projects apply for the next step of funding. The programme has consistently assessed both actor-level and 'system-level' innovation (narrowly defined, see Supplementary Appendix B)—tracking indicators of innovative capability (e.g. new partnerships, involvement of customers/users) and innovativeness linked to societal challenges (e.g. tangible solutions, willingness to invest). Different expectations are set for different stages of CDI projects, with the most ambitious expectations on longer-term system change in stage 3 projects (demonstrated by addressing multiple dimensions of systems change[11]). In addition, the external programme analysis provided formative evaluation and insights for future CDI calls. The evaluation strategy is based strongly on various ex-ante components including the programme logic model, selection criteria and CDI programme principles, which each consortium uses to draft its own specific project logic model when applying to CDI. The project-level logic model presents a clearly defined societal challenge whose solution addresses SDGs, outlines which innovations are targeted, and highlights the collaborative actions to be undertaken by the consortium of actors to drive system change. The implementation of these 'theories of change' is then followed in continuous monitoring activities. An ex-post programme analysis subsequently aims to capture and learn from system-level milestones reached by the projects.

4.1. 3 Strategic innovation programmes

The SIP evaluation strategy uses a blend of formative and summative approaches over the course of each programme's 12-year implementation period. The overall programme logic model and strategic innovation agenda for each SIP partnership establish directionality (ex-ante), which is monitored over time through annual reporting (of the portfolio of projects for each SIP) and reconciliation meetings between the funding agencies and SIP programme offices (focusing on drawing lessons from actions to date and identifying areas for further development). Comprehensive interim evaluations (conducted by external consultants) are conducted every 3 years, with slightly different foci for each. The first interim evaluation (after 3 years) is formative—assessing the establishment phase and shedding light on the programme's strengths and potential for improvement. The second interim evaluation (after 6 years) is the most comprehensive, including both summative (assessing results and early signs of longer-term effects from the SIP investment) and formative (related to programme strategy and organization) aspects. The 6-year interim evaluations also include a policy learning exercise to analyse how SIPs were contributing to key processes of sustainability transitions. The final interim evaluation (after 9 years) is primarily summative and includes recommendations on the closing phase (and exit strategy) for the SIP. Two key features of the SIP evaluation strategy are the multiple levels of analysis (projects, SIPs, meta-evaluation) pursued in the interim evaluations that would enable a nested approach to evaluation, as well as the experimental policy learning on system transitions that was conducted as part of the 6-year interim evaluation.

An overview of the evaluation strategy characteristics for all three cases is presented in Table 5 below.

4.2 Cross-case analysis

Based on the case analysis (see overview in Table 5), in this subsection, we analyse how well the evaluation strategy of

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

Table 5. Overview of evaluation strategies.

Dimensions	Categories	Vinnväxt	CDI	SIPs
Evaluation set up	Timing	• Ex-ante, robust recurring monitoring, interim evaluation (after 3 and 6 years), and ex-post	• Ex-ante, monitoring (status reports and limited strategic dialog), and ex-post	• Ex-ante (not formally done), monitoring (status reports and annual reconciliation meetings), interim evaluations (after 3, 6, and 9 years)
	Purpose	• Vinnväxt evaluation strategy involves different purposes • Recurring monitoring shows characteristics of summative, formative and developmental • Interim evaluations also focused on being formative (with some summative assessments) • Ex-post evaluations focused on summative and formative evaluation	• CDI evaluation strategy has primarily formative aims • Ex-ante evaluation has the purpose of describing the directionality (specific challenge) and input strengths (mobilized actors and potential) and expected results and effects (through project-level logic model) in relation to programme logic model • Monitoring focused on formative aims • Ex-post analyses focused on formative (programme level) and also summative (project level)	• Ex-ante evaluation has the purpose to describe how each SIP is associated with one or more SIAs, as well as explain expected results and impacts in relation to the programme's general impact goal. • Monitoring shows characteristics of formative evaluation, as it aims at improving programme planning and follow-up on the results with the funding authorities. • The interim evaluations vary between being formative and summative.
	Scope of the evaluation	• The scope of the evaluation is focused on the system of actors engaged in the initiative's regional area of strength • The evaluation is focused on renewal and strengthened system effectiveness (primarily STI policy), which can involve different instruments	• The scope of the evaluation is focused on the system of actors engaged in addressing the targeted challenge • The evaluation is mainly focused on development and scaling of a solution to address the challenge through innovation	• The scope of evaluation is on the system of actors engaged in the different innovation programmes—focused on areas of strategic importance for Sweden • The evaluation is focused on sustainable solutions to global societal challenges (primarily STI policy), and not entire policy mixes. These cover multiple instruments
	Level of analysis	• The level of analysis is the initiative's collection of projects, with limited analysis across different Vinnväxt initiatives at the programme level, i.e. no nesting approach	• The level of analysis is on individual projects • Ex-post analysis informs learning at the programme level	• Ex-ante evaluation is done at the level of programme • Recurring monitoring is done at the level of programme and projects • The interim evaluations focus on the evaluation of each programme (projects are considered at an aggregate level) • The SIPs evaluation strategy indicates a potential nested approach
	Stakeholder involvement	• The Vinnväxt evaluation strategy calls for a strong degree of stakeholder involvement • Recurring monitoring involves action researchers attached to each Vinnväxt initiative	• The CDI evaluation strategy encourages stakeholder involvement—particularly in the initial decisions on directionality/logic model and in the reflexive reporting at the end of each stage.	• The SIPs evaluation strategy does not cover a broad range of stakeholders, instead focusing mostly on participants and beneficiaries

(continued)

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

Table 5. (continued)

Dimensions	Categories	Vinnväxt	CDI	SIPs
Main questions	Input/output/outcomes	• Recurring monitoring (through annual reporting) collects both traditional input and output indicators (e.g. increased R&D investments, number of publications, number of innovations) and qualitative indicators of system-level milestones/outcomes (layer model) • External interim evaluations follow-up on signs of technological and business innovation • The ex-post self-assessment report highlights milestones (outcomes) over time • The evaluation results indicate that the programme has contributed to strengthening the regions' areas of specialization and the effectiveness of the innovation ecosystems (both in terms of innovation performance and contribution to system-level change)	• The consideration of input/output/outcomes are established with the project's logic model at the start (ex-ante) • Recurring monitoring focus on input and output indicators, by include some system-level effects • Ex-post surveys and self-assessment reports collect data on traditional innovation outputs and indicators of system-level and innovation capacity • The evaluation results reflect on the contribution of the projects to forming new actor networks and to addressing societal challenges	• The consideration of input/output/outcomes category vary depending on the timing of the evaluation. • Recurring monitoring focuses on the collection of statistics on the projects and rely mostly on traditional input and output indicators • The 6-year interim evaluation looks into how each SIP is contributing to transformative processes, based on the function of the innovation system and transition management processes. This, however, is done experimentally • The 6-year evaluation indicates that the SIPs have been contributing to knowledge development and resource mobilization in their targeted strategic areas. They have also contributed to collaboration across industries and value chains.
	Additionality	• Additionality is not formally addressed in the evaluation, although the layer model could be harnessed for this purpose in the recurring monitoring.	• Additionality is not formally captured in the analysis	• Only the 6- and 9-year interim evaluations cover additionality but to a limited extent
	Directionality	• Directionality is addressed in the evaluation strategy through the initiative's own logic model	• Directionality is addressed in the evaluation strategy through the project's own logic model (but the focus is on innovative solutions rather than at the level of socio-technical systems)	• Directionality is not addressed in the evaluation strategy
Methodology	Causality	• Mix of (primarily qualitative) methods to track system-level change, yet causality is not formally addressed in the evaluation strategy	• Mix of methods to track cumulative system-level change, yet causality is not formally addressed in the evaluation strategy	• Causality is not formally addressed

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

Table 6. Alignment between innovation programmes evaluation strategy with TIP evaluation.

Dimension	Categories	Programme alignment		
		Vinnväxt	CDI	SIPs
Evaluation set-up	Timing			
	Purpose			
	Scope of the evaluation			
	Level of analysis			
	Stakeholder involvement			
Main questions	Input/output/outcome			
	Additionality			
	Directionality			
Methodology	Causality (how to address causality)			

Inexistent/low, Medium, High.

each programme is aligned with the TIP evaluation dimensions introduced in Table 1, using the thresholds defined in Supplementary Appendix A. Table 6 shows the main findings across cases.

In the evaluation set-up dimension, the programmes use different *timings* of evaluation with different *purposes*. Vinnväxt has more formative and developmental purposes than the other programmes, so its evaluation strategy includes a variety of ex-ante, monitoring, interim and ex-post evaluation activities. Strong ex-ante and recurring monitoring components contribute to strong reflexive practices. CDI focuses on ex-ante evaluation via stage-based project selection criteria but also takes account of results ex-post, focusing less on monitoring. CDI analysis activities are primarily for learning and development, but also have some summative characteristics. Ex-ante assessment of the SIPs is confined to the proposal stage, and monitoring is done via traditional status reports. The interim evaluations in years 3, 6 and 9 track the SIPs through their long life-cycles, with the focus of evaluation evolving from a predominantly formative to a more summative style over the 9 years of the cycle.

The *scope* of the evaluations covers particular projects and programmes rather than inter-linkages with other programmes or contributions to other objectives beyond research and innovation policy. And although the scope of the programmes (particularly Vinnväxt and SIP) is on various types of systems and system-level change, none of the programme evaluations has the scope of broader socio-technical systems, instead focusing on geographic, challenge-driven or thematic innovation ecosystems.

The evaluation strategies work at different levels of analysis. Vinnväxt evaluation focuses on individual initiatives' project portfolios over time and does not compare across initiatives. The CDI evaluation strategy addresses the individual projects at different stages of implementation but does not group them by challenge or thematic areas. At the whole-programme level, there is interest in how well the CDI funding works. The SIPs seem to be on the way towards a more nested approach to evaluation. The overall SIP programme is a collection of individual SIPs, each with its own portfolio of projects. Its evaluation strategy covers these different levels at different times. Recurring monitoring is done at the level of projects and SIPs. The 3-yearly interim evaluations address the level of the individual SIPs. The overall SIP programme is tackled by meta-evaluations of the groups of individual SIPs started at the same time. In each case, the main weight of evaluation effort is at the level at which funding decisions are made, namely initiatives in Vinnväxt, projects in CDI, individual SIPs in the SIP programme.

All three programmes have some *stakeholder involvement* in the evaluation over and above participants, but not the more inclusive participation anticipated in the TIP literature. Vinnväxt is the most active in promoting broader stakeholder involvement, including not only programme leadership and advisory groups, but also action researchers 'attached' to each initiative (who, in turn, involve other participating stakeholders in reflexive processes). In CDI, both initiative participants and external experts are involved in the ex-ante evaluation. Programme monitoring is handled by the initiative management; ex-post evaluation is conducted by external consultants, with the involvement of initiatives' stakeholders and the programme team at Vinnova. The involvement of stakeholders is limited. The SIPs do not involve a broad range of stakeholders in the evaluation, which is centred around each SIP programme management and project participants.

For the 'main questions', the programmes use traditional *input* and *output* innovation indicators (e.g. R&D investments and patent applications) in traditional status reports and the interim evaluations of SIPs and Vinnväxt. The evaluation of *outcomes* differs, based on the timing of the evaluation and the specific conceptualization of the target system used by each programme (or for evaluation purposes). Vinnväxt uses a regional innovation system perspective and, during its monitoring activities, looks at systemic processes that resemble the TIS functions as part of its 'layer model' (see Supplementary Appendix B and Wise et al. 2022). CDI follows the definition of system innovation proposed by Miedzinski, Mazzucato and Ekins (2019) and assesses five dimensions of 'system innovation change' (again, not the same as transformative processes as defined in TIP) in a cumulative exercise that inform the selection of projects through the later stages of the stage-gate process. The SIPs do not have a common definition of system outcomes. However, the 6-year interim evaluation includes an experimental analysis of transformative processes (based on TIS functions and the transitions management literature), so systemic and transformative outcomes are considered to some degree.

Based on their system understanding, the evaluations also report on how the programmes contributed to systems change (using their specific views on systems). For Vinnväxt, the evaluators look at the contribution to the regions' areas of specializations as well as to the build-up of the triple helix model of innovation. In CDI, the evaluators report that individual projects have been able to develop knowledge and collaborations around specific solutions (not necessarily looking at the impact of these projects on innovation systems). For the SIPs, the evaluators highlighted that the programmes were able to develop knowledge and mobilize resources in their areas of strength. Additionally, they contributed to the build-up of new networks across industries and value chains.

(System-level) additionality is not formally addressed in the evaluation strategy of any of the programmes. While some additionality assessment is done in the SIPs interim evaluation, the process of assessment is not quantitative. From a TIP perspective, the programmes do not directly account for *directionality*. In CDI, directionality is assessed ex-ante in the stage-gate approach, where directionality is a criterion for selection. Both Vinnväxt's interim evaluation and CDI's stage-gate assessments allow some degree of re-orientation.

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

However, none of the programmes makes an assessment regarding the long-term direction of innovation and system change.

As to the methodology dimension, the programmes' evaluation strategies all include a variety of methods for data collection and analysis, but none of them formally accounts for *causality* by providing a deeper description and analysis of how the initiatives/projects/programmes contribute to changes in sociotechnical systems. This is perhaps because all of the programmes were conceived within the second generation of innovation policy—relying on linear logic models (where more innovation is better) instead of using theories of change that describe how the intervention is expected to lead to changes in the system which are viewed as necessary in order to address the targeted challenge.

5. Discussion

Our cross-case analysis shows that some aspects of TIP evaluation are already integral parts of Vinnova's existing programme evaluation practice (e.g. a robust set of activities to foster ongoing reflexive and participatory monitoring in Vinnväxt, ex-ante requirements for clearly defined directionality in CDI, and a system-wide scope and policy learning on transformative outcomes in SIPs). However, some cross-cutting challenges remain in becoming more aligned with TIP: (1) conceptualizing systems and outcomes, (2) developing transformative theories of change, and (3) addressing system-level additionality and directionality. While these challenges are evident in the findings from the cross-case analysis, they are not necessarily generalizable to other contextual settings. Therefore, the framework should be tested in other contexts o check whether such challenges are also identified elsewhere. Below, we discuss these three challenges and reflect on the implications of the analytical framework for TIP evaluation, while highlighting some opportunities for future work.

5.1 Conceptualizing systems and outcomes

The first challenge is related to the conceptualization of systems and outcomes used in the programmes. We argue that in order to have a more TIP-oriented evaluation strategy, evaluators need to consider the scope of the system in focus, as well as the type of system changes and outcomes the programme is targeting (which should be considered already during the design phase of the programme). Indeed, as highlighted by Molas-Gallart et al. (2021: 436), 'specific policies, their implementation, and evaluation should be coherent with the stated research and innovation policy objectives (directionality, societal goals, and system impact)'. As mentioned in Section 4.2, we found three different views on systems which influence and serve as bases for the evaluative strategies of the three Vinnova programmes and, hence, how they analyse outcomes. Vinnväxt follows a regional innovation system approach using seven categories of system effects (Wise et al. 2022). CDI follows the definition of system innovation by Miedzinski, Mazzucato and Ekins (2019), analysing how projects addressed different dimensions of system change. The SIPs seem to use a mix of the sectoral and the technological innovation system approaches. The evaluations, therefore, have reported on the outcomes based on these views of systems, i.e. on the build-up of regional (Vinnväxt) and technological/sectoral innovation systems (SIPs) and the contribution of these collaborative efforts to broader system-level change processes, and the contribution of projects to some system-level aspects, narrowly defined (CDI).

While using different understandings of systems is not problematic, adapting programmes to be more transformative means having a broader perspective on systems and the outcomes they aim to achieve. Here we apply the definition of systems used in the sustainability transitions literature, focusing on the transition of entire sectoral socio-technical systems of production and consumption (e.g. the energy or transport sectors), where outcomes are understood as innovation and transition processes that enable change in sociotechnical systems (Luederitz et al. 2017; Molas-Gallart et al. 2021; Haddad and Bergek 2023) in order to fulfil key societal functions (see Geels 2004), as discussed in Section 2.2. This definition of system innovation seems to be widely accepted in the TIP literature (Schot and Steinmueller 2018; Diercks, Larsen and Steward 2019; Haddad et al. 2022).

Another branch of TIP research, the mission-oriented innovation policy (MOIP) approach, can use a different understanding of the system(s) targeted by TIP. For example, Miedzinski, Mazzucato and Ekins (2019) define system innovation as a collection of innovations that have transformative potential. In that case, the focus is not on socio-technical systems (or innovation systems) as such, and it is not clear what the target is. The authors list a few dimensions of change related to their view on system innovation, including culture and values, regulatory framework, infrastructure and production systems, business models, technologies, products, and processes. However, these elements are simplifications of some of the processes that could be analysed and are not as comprehensive as, for example, the TIS functions. Other researchers seem to understand a third-generation MOIP as a way to re-orient innovation systems (technological, sectoral or national) towards addressing societal missions based on grand challenges (Mazzucato 2016; Kattel and Mazzucato 2018; Mazzucato 2018). These papers, however, focus more on the general characteristics of missions rather than on evaluation, and they do not propose any specific framework (such as the transformative outcomes or TIS functions) for assessing mission outcomes.

More recently and similar to other TIP scholars (e.g. Ghosh et al. 2021; Bergek and Haddad 2022), Elzinga et al. (2023) proposed using the innovation system perspective as a basis to assess mission progress. As mentioned in Section 2.2, the authors propose a set of (9) functions divided between programming (e.g. directionality and coordination) and performance functions (related to processes such as knowledge development and entrepreneurial experimentation, as well as processes related to destabilization).

Our point here is that policymakers and evaluators should be careful when delineating the type of systems and the scope of systems change processes they are targeting, especially if the goal is to make programmes (and their evaluations) more TIP-oriented. The type of outcomes evaluators should be looking for in TIP are those processes contributing to sociotechnical system change. In this regard, while there is no problem in working with different conceptualizations of system outcomes, it is helpful to be clear which is being used in order to adopt a consistent approach.

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

5.2 Developing (transformative) theories of change

The second challenge is to develop transformative ToCs. Within this challenge, two main issues can be highlighted. We observe that all three programmes' evaluations rely mostly on logic models rather than complete theories of change, which would specify causal mechanisms and contextual influences, as well as consider unintended effects. The logic models found in the three Swedish programmes provide overviews of the logic behind the programme by highlighting some of the expected outputs and outcomes of the programmes, but do not specify causal mechanisms. Mayne (2015) would call this an 'overview ToC', which represents just an intermediary version of the ToC, which should be complemented by a causal ToC showing the causal link assumptions. When dealing with more complex programmes, Funnell and Rogers (2011) suggest some alternatives to portray complex ToCs. These include using diagrams from system approaches, such as system dynamic models, or developing an emergent theory of change, which should be adapted during programme implementation. In the three Swedish programmes considered, however, the logic models simply describe in general terms inputs, activities, outputs, and outcomes and do not consider causal mechanisms.

Other insights on how to design more transformative ToCs have also been put forward in the TIP literature. For example, Haddad and Bergek (2023) propose a framework that incorporates the development of a programme theory for TIP interventions based on the combination of the realist approach and ToC. Following the authors' suggestion, the ToC is complemented by the identification of causal mechanisms, underpinned by the transformative processes, such as those proposed by Bergek and Haddad (2022) and Ghosh et al. (2021) (see Section 2.2). TIPC has also developed a handbook on how to design transformative theories of change (Boni et al. 2021), also relying on the transformative outcomes proposed by Ghosh et al. (2021).

An additional issue that will need to be taken up when designing more transformative ToCs is the need for them to reflect contextual influences and the framing (or boundaries) of the system changes that are targeted. This would include acknowledging that innovation programmes are built around specific sectors and political and economic contexts, as well as being influenced by context structures and landscape pressures. This is because most of the current innovation programmes are path-dependent and based on policy layering. Rohracher, Coenen and Kordas (2023) specifically look at how the SIPs have been based on layering practices. This makes contextual influences particularly relevant when designing more transformative ToCs.

5.3 Addressing system-level additionality and directionality

The third shortcoming relates to the assessment of system-level additionality and directionality Our findings show that these two categories, which belong to the 'main questions' dimension of the framework, are addressed to a limited extent in the evaluation strategies of the three programmes. Regarding additionality, the need to address behavioural additionality has already been discussed extensively in the literature, but evaluations still fail to capture it at the system level (Gök and Edler 2012; Amanatidou et al. 2014). Our analysis showed that only the SIPs considered additionality, but to a limited extent, as this was focused based on participants' perception of how the programme contributed to bringing additional inputs and actor-network effects. According to the literature, covering system-level additionality, in contrast, would require going beyond these aspects to understand the contribution of the programmes to transformative processes (Haddad and Bergek 2023). For example, evaluators could investigate how the programmes contributed to changes in actor networks, in terms of entry of niche actors and reorientation of incumbent actors towards support transitions, or the impact it had in entrepreneurial experimentation and market formation processes. This could be done, for example, by unpacking the causal mechanisms at play using process tracing methods. The transformative ToC and the type of system change the programme is targeting, previously identified when scoping the evaluation (see Section 5.1), can also serve as a baseline to inform additionality assessment. The TIP literature provides more guidance on how to address additionality at the system level, as discussed in Section 2.2 (Janssen 2019; Haddad and Bergek 2023). However, there is a lack of empirical applications operationalizing and illustrating the approaches proposed by these authors, which should be addressed in future research.

The cross-case analysis shows that directionality seems to be considered more in relation to ex-ante evaluations than in later evaluation stages (although it to some extent contributes to the reflexivity of the Vinnväxt and CDI programmes during their implementation). Given the long-term nature of sustainability transitions, considering directionality in the evaluation should also be focused on assessing the direction of change the programmes are enabling. Even if some of these changes might only be seen (long) after specific programmes are finished, evaluations should be able to capture early signs of change. As discussed in Section 2.2, some approaches can be used in order to do this, e.g. assessing transition pathways (Turnheim et al. 2015; Geels et al. 2016). However, to the best of our knowledge, no evaluation to date has tried to apply any of these approaches.

It is worth noting that the operationalization of additionality and directionality in any TIP evaluation strategy is also connected to how interventions define systems, as well as how clear the ToC is, given that it specifies how the programme is supposed to work and contribute to transformative change.

5.4 Implications of the analytical framework for TIP evaluation

As we mentioned in Section 2, we adapt the framework proposed by Edler et al. (2012) to develop an analytical framework to identify the main characteristics of an archetypical evaluation strategy for TIP. The framework allowed us to identify several key issues underlying the Swedish innovation programmes and to reflect how they align with this archetypical model, in case they were to be designed for TIP purposes. We acknowledge that the dimensions portrayed in the framework are non-exhaustive, and additional dimensions might be needed in the future. For example, the framework could reflect the evaluation of policy mixes. This is particularly key in TIP evaluation, as the interaction of policy mixes is relevant for transformative change. Rogge and Reichardt (2016), for example, identify three building blocks of policy mixes: (1) they are composed of elements such as policy strategies and instrument mixes, (2) these elements are results of policy

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

processes, (3) both elements and processes can be described by their characteristics such as coherence, credibility, and comprehensiveness. All these building blocks can have implications for TIP evaluation and should be considered in future work. This is especially important given that, while STI policy has taken up the need to transform entire sociotechnical systems to address societal challenges, doing so will need to go beyond the remit of STI policy (OECD 2023). Therefore, STI policies will need to be combined with policies from other domains, such as environment and climate.

Another dimension that might need to be considered is policy coordination. This is not only related to the coordination of policy mixes, but also to how policy is implemented around different levels of government, i.e. local, regional, national and supranational, but also within different agencies and ministries (Arnold and Barker 2022). This will also bring additional complexity to the evaluation. Coordination is also particularly relevant when considering context in evaluation, given that governments' structures and operations, as well as political systems, impact how programmes are designed and implemented.

Including these categories will also affect other dimensions of the framework. For example, evaluating entire policy mixes and coordination will also have consequences for the scope of the evaluation. Among other things, the scope will need to be broadened to account for new instruments beyond STI policy, their interactions, and how they affect and are affected by changes in socio-technical systems. Moreover, the pool of stakeholders will be bigger, e.g. due to the involvement of different ministries, agencies and other actors. Further, contextual factors will become even more critical, given that factors such as path dependencies among ministries and industries have the potential to affect policy implementation and their capacity to promote change. This will also put great demand on ToCs, which will need to be adapted to reflect such interdependencies and complexity.

6. Conclusions

This paper looks at the theory and practice of TIP evaluation and makes three main contributions, with implications for policy and programme design, evaluation practice and future knowledge development. First, we propose an analytical framework describing the main dimensions and characteristics of an archetypal TIP evaluation strategy. Specifically, we tweak Edler et al.'s framework and reflect how evaluation dimensions such as evaluation set-up, main questions and methodology differ for TIP. The changed characteristics in the proposed analytical framework for TIP evaluation could potentially have implications on programme design (where planning of evaluation strategies begins), and on evaluation practice (where methodological approaches need to be adapted to be more TIP-oriented).

Second, in an initial test of the framework, we show the extent to which existing evaluation practices of three Swedish innovation policy programmes, i.e. Vinnväxt, CDI and SIPs are aligned with an archetypal TIP evaluation strategy. These programmes have been conceived within the second generation of innovation policy, hence their approaches to evaluation are not fully aligned with an archetypal TIP evaluation strategy. Nonetheless, our findings show advances in some categories. These include considerations of directionality in CDI, an evaluation strategy that emphasizes learning and reflexivity in Vinnväxt, and the experimental assessment of more transformative-oriented processes in the SIPs. Additionally, while our cross-case analysis serves as a 'diagnostic' of the alignment of Vinnova's programmes, it can potentially function as an example and learning exercise for policymakers and practitioners beyond Sweden and shed light on key aspects to consider when designing TIP programme evaluation strategies.

Finally, we identify three challenges involved in adapting current innovation policy evaluation practice to TIP that policymakers and practitioners should consider and outline potential ways to deal with these challenges. These are: (1) the consideration of the type of system the programmes aim to change and the outcomes the programme is targeting, given that TIP demands a broader scope of evaluation to consider changes in entire socio-technical systems; (2) the need to develop ToCs that reflect and unfold key transformative outcomes (as processes), elaborating causal mechanisms and contextual influences; and (3) the need to find new and practical ways to track system-level additionality and directionality over time in order to allow for learning and adaption of programmes. We also reflect on the key implications of the framework and how it can be further improved.

This paper takes first steps in defining an analytical framework for TIP evaluation and outlining key dimensions and methodological considerations that can inform policymakers and practitioners in their ongoing adaptations to TIP. Future research can support these efforts through additional work to address the identified challenges and through further development and testing of the framework for TIP evaluation in other settings.

Acknowledgements

The authors would like to thank Professor Anna Bergek for her guidance and helpful comments. We also appreciate the comments from Philippe Larrue and other participants at the Eu-SPRI 2023, 14–16 June 2023, Brighton, UK. We would like to express our sincere gratitude to the anonymous reviewers for their invaluable feedback and constructive criticism, which significantly contributed to enhancing the quality and clarity of this manuscript.

Supplementary data

Supplementary data are available at *Research Evaluation Journal* online.

Funding

This work was supported by funding from the Swedish Innovation Agency Vinnova through the Governing Resilience and Transformation (GReaTr) project [grant number: 2021-02077]. The funder has not been involved in study design; data collection, analysis, or interpretation; the writing of the article; or the decision to submit the article for publication.

Notes

1. Evaluation strategy is a term used by policy practitioners to describe the set of activities used in relation to monitoring, evaluation and learning (MEL) from implementation of policy interventions (see e.g.

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

OECD 2020). We use this term throughout the paper to similarly refer to the set of monitoring, evaluation and learning (MEL) activities undertaken over time (ex-ante, ex-durante, interim and ex-post).
2. Edler et al. (2012) propose a fifth category, i.e. policy and quality, which covers the usefulness of policy recommendations, consequences and quality of the evaluation. We, however, do not include this dimension given that this is not expected to differ between TIP and traditional approaches to innovation policy evaluation.
3. For example, this refers to the ASIRPA$^{\text{Real-time}}$, a framework and toolset to help project and programme managers to guide activities towards societal goals, based on an iterative and formative approach (Matt et al. 2023).
4. The work on policy mixes has also highlighted the 'top-down' versus 'bottom-up' approaches for assessing policy mixes, which can also shed light on the scope of the evaluation (Rogge and Stadler 2023). The former aims at capturing the intentional policy mix design and would start by specifying the overarching policy strategy intent and the combination of instruments used to achieve it. In this approach, strategies and instruments adopted in other policy domains and governance levels would not be considered part of the assessment. In contrast, the latter approach would look for all instruments influencing a particular field, independently if they were designed with that purpose from the outset or if they are coming from other policy domains.
5. The TIS functions include processes related to knowledge development and diffusion, entrepreneurial experimentation, resource mobilisation, guidance of search, market formation, legitimation, and development of positive externalities (Hekkert et al. 2007; Bergek et al. 2008; Bergek 2019).
6. In contrast, previous approaches to science and technology policy, and innovation systems, focus on changes at the firm level, and on spatially (regional and national) or cognitively (technological and sectoral) delineated innovation systems (Chaminade and Edquist 2010).
7. Specifically, the authors widen the TIS functions to cover both innovation and transition-related processes, by looking at the socio-technical transitions literature (i.e. MLP and Strategic Niche Management (SNM)). The authors argue that, while the TIS functions already cover most of the processes explaining changes in socio-technical systems, they give less attention to structural dynamics related to changes in actor networks and institutions. As a result, the authors propose three sets of transformative processes to capture changes in: (1) socio-technical system, (2) actor networks, and (3) institutions (Bergek and Haddad 2022).
8. For a more detailed description of each type of causality, see Geels (2022). In short, conjunctural causality explains outcomes as the result of the convergence and alignment of independent events and processes in a certain moment in time. Configurational causality explains the way in which multiple interacting components of large-scale, heterogeneous entities combine or co-evolve to produce different outcomes. Event-chain causality explains how sequential interactions between events generate outcomes over time.
9. The 10-year funding is divided into three phases. After the initial approved application, decisions to continue funding Vinnväxt initiatives (or not) are based on results from interim evaluations after 3 and 6 years. After the three phases of funding, a final ex-post assessment of effects is made.
10. The SIPs were funded by Formas and the Swedish Energy Agency, in addition to Vinnova.
11. Using five dimensions of system innovation from Miedzinski, Mazzucato and Ekins (2019).

References

Alkin, M. C. (2013) *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences*. Thousand Oaks, US: SAGE Publications.

Amanatidou, E. et al. (2014) 'Using Evaluation Research as a Means for Policy Analysis in a 'New' Mission-Oriented Policy Context', *Minerva*, 52: 419–38.

Arnold, E. (2004) 'Evaluating Research and Innovation Policy: A Systems World Needs Systems Evaluations', *Research Evaluation*, 13: 3–17.

Arnold, E., and Barker, K. E. (2022) 'What Past changes in Swedish Policy tell us about Developing Third-Generation Research and Innovation Governance', in M. Benner, G. Marklund and S. Schwaag Serger (eds.) *Smart Policies for Societies in Transition: The Innovation Challenge of Inclusion, Resilience and Sustainability*, pp. 59–86. Cheltenham: Edward Elgar Publishing.

Arnold, E. et al. (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, accessed 06 Jul. 2023.

Arnold, E., and Wise, E. (Forthcoming) 'Evaluating Transitions and Missions: It's All About the Theory', in: U. Cantner, D. Fornahl, and S. Kuhlmann (eds) *The New Role of the State for Transformative Innovation*. Cheltenham, UK: Edward Elgar.

Åström, T., and Arnold, E. (2023) 'Meta-Evalaution of the Swedish Strategic Innovation Programmes: As Policy Instruments for Industrial Competitiveness and Systems Innovation', Vinnova Report VR 2023:16. downloaded on 11 March 2024 from https://www.vinnova.se/publikationer/meta-evaluation-of-the-swedish-strategic-innovation-programmes/, accessed 05 Jul. 2024.

Bartle, D., and Morris, M. (2010) 'Evaluating the Impacts of Government Business Assistance Programmes: Approaches to Testing Additionality', *Research Evaluation*, 19: 275–80.

Bergek, A. (2019) 'Technological Innovation Systems: A Review of Recent Findings and Suggestions for Future Research', in F. Boons and A. McMeekin (eds) *Handbook of Sustainable Innovation*, pp. 200–218. Cheltenham, UK: Edvard Elgar Publishing.

Bergek, A. et al. (2008) 'Analyzing the Functional Dynamics of Technological Innovation Systems: A Scheme of Analysis', *Research Policy*, 37: 407–29.

Bergek, A., and Haddad, C. R. (2022) 'Evaluating Transformative Innovation Policy Outcomes as Unfolding Processes of Change in Sociotechnical Configurations', in G. Vélez-Cuartas, and O.Y. Romero-Goyeneche (eds) *Transformative Metrics: Contributions to the Studies for Monitoring and Evaluating How Science, Technology, and Innovation Can Address Social and Environmental Callenges*, pp. 63–105. Medellín, CO: Universidad de Antioquia.

Bergek, A., Hellsmark, H., and Karltorp, K. (2023) 'Directionality Challenges for Transformative Innovation Policy: Lessons from Implementing Climate Goals in the Process Industry', *Industry and Innovation*, 30: 1110–39.

Boni, A. et al. (2023) 'Evaluating Transformative Innovation Policy in a Formative Way: Insights from Vinnova's Food Mission Experiment', *Research Evaluation*, 32: 577–90.

Boni, A. et al. (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/, accessed 07 Jul. 2023.

Borrás, S., and Laatsit, M. (2019) 'Towards System Oriented Innovation Policy Evaluation? Evidence from EU28 Member States', *Research Policy*, 48: 312–21.

Borrás, S., and 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: 659–72.

Bos, M. W., Hofman, E., and Kuhlmann, S. (2016) 'An Assessment Method for System Innovation and Transition (AMSIT)', *International Journal of Foresight and Innovation Policy*, 11: 185–214.

Bowen, G. A. (2009) 'Document Analysis as a Qualitative Research Method', *Qualitative Research Journal*, 9: 27–40.

Bryman, A., and Bell, E. (2015) *Business Research Methods*, 4th edn. Oxford, UK: Oxford University Press.

Chaminade, C., and Edquist, C. (2010) 'Rationales for Public Policy Intervention in the Innovation Process: A Systems of Innovation Approach', in R.E. Smits, S. Kuhlmann, and P. Shapira (eds) *The Theory and Practice of Innovation Policy: An International Research Handbook*, pp. 95-114. Cheltenham: Edward Elgar Publishing.

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

Clarysse, B., Wright, M., and Mustar, P. (2009) 'Behavioural Additionality of R&D Subsidies: A Learning Perspective', *Research Policy*, 38: 1517–33.

de Oliveira, L. G. S., Lacerda, J. S., and Negro, S. O. (2020) 'A Mechanism-Based Explanation for Blocking Mechanisms in Technological Innovation Systems', *Environmental Innovation and Societal Transitions*, 37: 18–38.

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., and Steward, F. (2019) 'Transformative Innovation Policy: Addressing Variety in an Emerging Policy Paradigm', *Research Policy*, 48: 880–94.

Edler, J. et al. (2012) 'The Practice of Evaluation in Innovation Policy in Europe', *Research Evaluation*, 21: 167–82.

Edler, J., and Boon, W. P. (2018) 'The Next Generation of Innovation Policy: Directionality and the Role of Demand-Oriented Instruments'—Introduction to the Special Section', *Science and Public Policy*, 45: 433–4.

Eisenhardt, M. E. (1989) 'Building Theories from Case Study Research', *The Academy of Management Review*, 14: 532–50.

Elzinga, R. et al. (2023) 'Assessing Mission-Specific Innovation Systems: Towards an Analytical Framework', *Environmental Innovation and Societal Transitions*, 48: 100745.

Flanagan, K., Uyarra, E., and Laranja, M. (2011) 'Reconceptualising the 'Policy Mix'for Innovation', *Research Policy*, 40: 702–13.

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

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: 1–14.

Geels, F. W. et al. (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.

George, A. L., and Bennett, A. (2004) *Case Studies and Theory Development in the Social Sciences*. Cambridge, USA: MIT Press.

Georghiou, L., and 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, accessed 07 Jul. 2023.

Ghosh, B. et al. (2021) 'Transformative Outcomes: Assessing and Reorienting Experimentation with Transformative Innovation Policy', *Science and Public Policy*, 48: 739–56.

Gök, A., and Edler, J. (2012) 'The Use of Behavioural Additionality Evaluation in Innovation Policy Making', *Research Evaluation*, 21: 306–18.

Grillitsch, M. et al. (2019) 'Innovation Policy for System-Wide Transformation: The Case of Strategic Innovation Programmes (SIPs) in Sweden', *Research Policy*, 48: 1048–61.

Haddad, C. R. et al. (2022) 'Transformative Innovation Policy: A Systematic Review', *Environmental Innovation and Societal Transitions*, 43: 14–40.

Haddad, C. R., and Bergek, A. (2023) 'Towards an Integrated Framework for Evaluating Transformative Innovation Policy', *Research Policy*, 52: 104676.

Hekkert, M. P. et al. (2007) 'Functions of Innovation Systems: A New Approach for Analysing Technological Change', *Technological Forecasting and Social Change*, 74: 413–32.

Hind, J. (2010) 'Additionality: A Useful Way to Construct the Counterfactual Qualitatively?', *Evaluation Journal of Australasia*, 10: 28–35.

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., and Wesseling, J. H. (2022) 'Evaluating Systemic Innovation and Transition Programmes: Towards a Culture of Learning', *PLOS Sustainability and Transformation*, 1: e0000008.

Kattel, R., and Mazzucato, M. (2018) 'Mission-Oriented Innovation Policy and Dynamic Capabilities in the Public Sector', *Industrial and Corporate Change*, 27: 787–801.

Kern, F. (2012) 'Using the Multi-Level Perspective on Socio-Technical Transitions to Assess Innovation Policy', *Technological Forecasting and Social Change*, 79: 298–310.

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

Kivimaa, P., and Kern, F. (2016) 'Creative Destruction or Mere Niche Support? Innovation Policy Mixes for Sustainability Transitions', *Research Policy*, 45: 205–17.

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

Kontigo (2016) 'Effektanalys av Vinnväxt-programmet: Analys av effekter och nytta (Effect Analysis of the Vinnväxt Programme: Analysis of Effects and Benefits', downloaded on 13 March 2023 from https://www.vinnova.se/contentassets/b1b047ac317b472c8547f7637727be0a/va_16_03t.pdf?cb=20170620161510, accessed 07 Jul. 2023.

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, accessed 07 Jul. 2023.

Luederitz, C. et al. (2017) 'Learning through Evaluation—A Tentative Evaluative Scheme for Sustainability Transition Experiments', *Journal of Cleaner Production*, 169: 61–76.

Magro, E., and Wilson, J. R. (2013) 'Complex Innovation Policy Systems: Towards an Evaluation Mix', *Research Policy*, 42: 1647–56.

Magro, E., and Wilson, J. R. (2019) 'Policy-Mix Evaluation: Governance Challenges from New Place-Based Innovation Policies', *Research Policy*, 48: 103612.

Matt, M. et al. (2023) 'ASIRPA$^{\text{Real-Time}}$ in the Making or How to Empower Researchers to Steer Research Towards Desired Societal Goals', *Research Evaluation*, 32: 412–25.

Mayne, J. (2015) 'Useful Theory of Change Models', *Canadian Journal of Program Evaluation*, 30: 119–42.

Mazzucato, M. (2016) 'From Market Fixing to Market-Creating: A New Framework for Innovation Policy', *Industry and Innovation*, 23: 140–56.

Mazzucato, M. (2018) 'Mission-Oriented Innovation Policies: Challenges and Opportunities', *Industrial and Corporate Change*, 27: 803–15.

Miedzinski, M., Mazzucato, M., and 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, accessed 07 Jul. 2023.

Molas-Gallart, J. et al. (2021) 'A Formative Approach to the Evaluation of Transformative Innovation Policies', *Research Evaluation*, 30: 431–42.

Molas-Gallart, J., and Davies, A. (2006) 'Toward Theory-Led Evaluation: The Experience of European Science, Technology, and Innovation Policies', *American Journal of Evaluation*, 27: 64–82.

OECD (2016) *OECD Review of Innovation Policy: Sweden 2016*. Paris: OECD Publishing.

OECD (2020) 'Improving Governance with Policy Evaluation: Lessons From Country Experiences', downloaded on 12 May 2023 from https://www.oecd.org/gov/improving-governance-with-policy-evaluation-89b1577d-en.htm, accessed 06 Jul. 2023.

OECD (2023) *Science, Technology and Innovation Outlook 2023: Enabling Transitions in Times of Disruption*. Paris: OECD Publishing.

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

Patton, M. Q. (2010) *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York, USA: The Guilford Press.
Pawson, R., and Tilley, N. (1997) *Realist Evaluation*. London, UK: SAGE Publications.
Ramboll (2020) 'Analys av Programmet Utmaningsdriven Innovation: Slutrapport (Analysis of the Challenge-Driven Innovation Programme: Final Report', downloaded on 13 March 2023 from https://www.vinnova.se/publikationer/udi—utvardering/, accessed 04 Jul. 2023.
Rogge, K. S., and Reichardt, K. (2016) 'Policy Mixes for Sustainability Transitions: An Extended Concept and Framework for Analysis', *Research Policy*, 45: 1620–35.
Rogge, K. S., and Stadler, M. (2023) 'Applying Policy Mix Thinking to Social Innovation: From Experimentation to Socio-Technical Change', *Environmental Innovation and Societal Transitions*, 47: 100723.
Rohracher, H., Coenen, L., and Kordas, O. (2023) 'Mission Incomplete: Layered Practices of Monitoring and Evaluation in Swedish Transformative Innovation Policy', *Science and Public Policy*, 50: 336–49. https://doi.org/10.1093/scipol/scac071.
Schot, J., and Steinmueller, W. E. (2018) 'Three Frames for Innovation Policy: R&D, Systems of Innovation and Transformative Change', *Research Policy*, 47: 1554–67.
Schwaag Serger, S., and Palmberg, C. (2022) 'Towards Transformative Policy in Finland and Sweden: Some Viewpoints from Practice', in M. Benner, G. Marklund, and S. Schwaag Serger (eds) *Smart Policies for Societies in Transition The Innovation Challenge of Inclusion, Resilience and Sustainability*, pp. 143-188. Cheltenham, UK: Edward Elgar Publishing.
SIPER (2023) 'Welcome to SIPER', https://si-per.eu/siper-en/, accessed 07 Jul. 2023.
Stufflebeam, D. (2001) 'Evaluation Models', *New Directions for Evaluation*, 2001: 7–98. https://doi.org/10.1002/ev.3.
Turnheim, B. et al. (2015) 'Evaluating Sustainability Transitions Pathways: Bridging Analytical Approaches to Address Governance Challenges', *Global Environmental Change*, 35: 239–53.
Ulmanen, J., Bergek, A., and Hellsmark, H. (2022) 'Lost in Translation: Challenges in Creating New Transformative Innovation Policy Practices', *PLoS Sustainability and Transformation*, 1: e0000031.
Velasco, D., and Witte, J. (2022) *An Analysis of Vinnova's Mission to Transform the Swedish Food System*. https://www.tipconsortium.net/an-analysis-of-vinnovas-mission-to-transform-the-swedish-food-system/, accessed 06 Apr. 2023.
Vinnova (2021a) https://www.vinnova.se/m/ekosystem-for-innovativa-foretag/vinnvaxt/om-vinnvaxt/, accessed 11 May 2023.
Vinnova (2021b) 'Avropsförfrågan SIP nioårsutvärdering', Call number 2021-02735, Stockholm: Vinnova.
Vinnova (2022a) *Impact Innovation—mobilisera för hållbar utveckling*. https://www.vinnova.se/m/impact-innovation/, accessed 31 Jan. 2023.
Vinnova (2022b) *Strategiska innovationsprogram: Samarbete för hållbar innovation*. https://www.vinnova.se/m/strategiska-innovationsprogram/, accessed 11 May 2023.
Vinnova (2022c) *Strategiska innovationsprogram: Samarbete för hållbar innovation [Strategic Innovation Programs: Collaboration for Sustainable Innovation]*. https://www.vinnova.se/m/strategiska-innovationsprogram/, accessed 19 Jan. 2023.
Wanzenböck, I. et al. (2020) 'A Framework for Mission-Oriented Innovation Policy: Alternative Pathways through the Problem–Solution Space', *Science and Public Policy*, 47: 474–89.
Weber, K. M., and 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: 1037–47.
Weiss, C. H. (1995) 'Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families', in J.P. Connell, A.C. Kubisch, L.B. Schorr, and C.H. Weiss (eds) *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*, pp. 65–92. Washington, USA: The Aspen Institute.
White, H. (2010) 'A Contribution to Current Debates in Impact Evaluation', *Evaluation*, 16: 153–64.
Williams, S., and Robinson, J. (2020) 'Measuring Sustainability: An Evaluation Framework for Sustainability Transition Experiments', *Environmental Science & Policy*, 103: 58–66.
Wise, E. et al. (2022) 'A Participatory Approach to Tracking System Transformation in Clusters and Innovation Ecosystems—Evolving Practice in Sweden's Vinnväxt Programme', *Research Evaluation*, 31: 271–87.
Wise, E., and Arnold, E. (2022) 'Evaluating Transformation—What Can We Learn from the Literature?', Papers in Innovation Studies no. 2022/10. Lund University.

Downloaded from https://academic.oup.com/rev/article/doi/10.1093/reseval/rvae059/7924330 by guest on 09 January 2025

© The Author(s) 2024. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Research Evaluation, 2024, 33, 1–17
https://doi.org/10.1093/reseval/rvae059
Article