



LUND
UNIVERSITY

CIRCLE

**CENTRE FOR
INNOVATION RESEARCH**

Evaluating Transformation – what can we learn from the literature?

Emily Wise, emily.wise@fsi.lu.se

CIRCLE & LU Collaboration Office, Lund University, Sweden

Erik Arnold, erik.arnold@technopolis-group.com

Manchester Institute of Innovation Research, University of Manchester, UK

Technopolis Ltd, Brighton, UK

Papers in Innovation Studies no. 2022/10

Want to stay informed about new working papers? Follow us on Twitter http://twitter.com/circle_lu or visit us on <http://www.circle.lu.se/publications/>.

The Papers in Innovation Studies series is open to all researchers working on innovation. Interested in submitting a working paper? Contact the editor: torben.schubert@circle.lu.se

The authors ensure that they own the copyrights of this material or have obtained the permission to publish it in this series from the copyright owners.

Evaluating Transformation – what can we learn from the literature?

Emily Wise*¹ and Erik Arnold²

Abstract

The last decade's rise of the so-called "third frame" (or third generation) of Transformative Innovation Policies (TIP) has shifted focus of research and innovation investments from economic growth and competitiveness to also tackling societal challenges and generating broader environmental and societal impact. The evolution in rationale and aims for policy action also implies a need to adapt and evolve evaluative strategies and practices. As policymakers begin to develop new transformative innovation programmes, a key question arises as to how monitoring, evaluation and learning practices (currently framed around 1st and 2nd generation innovation policies) can be adapted in order to meet 3rd generation innovation policy needs? In shaping a response, one can learn from both theory and practice. This brief (produced within the GReaTr initiative³) aims to provide a synthesis of what recent academic research tells us about evaluating transformation, leveraging a set of 11 seminal articles (and other complementary literature) to answer four questions: For whom and why? What to evaluate? How to evaluate? What unit of analysis?

The synthesis points to a relative consensus in the academic literature on the main purposes and uses, the recommended principles and approaches, as well as possibilities for delineating and dealing with multiple scopes and units of analysis in evaluating transformation – yet highlights different conceptual framings of system change. The summary provides inputs to planning an evaluative strategy for TIP and highlights the need to consider new questions related to approaches to reporting across funding agencies, and more active roles for funding and policymaking agencies in dialogues about strategic direction and prioritisation of investments.

Keywords: Innovation policy, Evaluation, System evaluation, Transformation, Socio-technical transitions

JEL code: O32, O38

* Corresponding author

¹ Centre for Innovation Research (CIRCLE) and LU Collaboration Office, Lund University, Medicon Village, Scheeleorget 1, 221 00 Lund, Sweden, emily.wise@fsi.lu.se

² Manchester Institute of Innovation Research, Alliance Manchester Business School, University of Manchester and Technopolis Ltd., 3 Pavilion Buildings, Brighton, BN1 1EE, United Kingdom, erik.arnold@technopolis-group.com

³ "Governing resilience and transformation – GReaTr" is a strategic initiative (co-funded by Vinnova and based at Lund University) which aims to function as a forum for interactive knowledge generation and use by providing a 'thinking space for policy' (a hybrid between research, analysis and advice), developing knowledge (e.g. policy briefs, etc.) and supporting its application in policy practice.

Summary

The last decade's rise of the so-called "third frame" (or third generation) of Transformative Innovation Policies (TIP) (Weber and Rohracher 2012; Schot and Steinmueller 2018) has shifted focus of research and innovation investments from economic growth and competitiveness to also tackling societal challenges and generating broader environmental and societal impact. The evolution in rationale and aims for policy action also implies a need to adapt and evolve evaluative strategies⁴ and practices.

TIP evaluation needs to account for different characteristics of policy implementation (e.g. longer-term and broader system-level change, coordination of action across multi-layers of governance, integration of participatory processes for reflection and feedback loops, flexible strategic reorientation). As policymakers begin to develop new transformative innovation programmes, a key question arises: **How can monitoring, evaluation and learning practices** (currently framed around 1st and 2nd generation innovation policies) **be adapted in order to meet 3rd generation innovation policy needs?**

In shaping a response, one can learn from both theory and practice. This first GReaTr policy brief aims to provide a synthesis of what recent academic research tells us about evaluating transformation, leveraging a set of 11 seminal articles (and other complementary literature) to answer four questions:

For whom and why?

- Who are the customers?
- What are the main purposes and uses of evaluation?

What to evaluate?

- What is TIP designed to do?
- What aspects of system transformation are important to evaluate?

How to evaluate?

- What recommended principles and approaches?
- Which elements of an evaluative strategy?

What unit of analysis?

- How to scope/delineate the boundaries of LT transformative processes?
- How to deal with MLG (multiple scopes)?

The synthesis points to a relative consensus in the academic literature on the main purposes and uses, the recommended principles and approaches, as well as possibilities for delineating and dealing with multiple scopes and units of analysis in evaluating transformation. Although presented in a variety of conceptual framings, the academic literature highlights four key aspects of system change/transformation that should be assessed in TIP evaluation: the goal setting (directionality and strategic orientation), the system elements that are changing, the progression (depth/scope) of system change, and the momentum provided through agency and transformative capacity.

While the summary overview provides helpful inputs to planning an evaluative strategy for TIP, it also highlights the need to consider new questions related to approaches to data collection/reporting across funding agencies, and different (more active) roles for funding and policymaking agencies in dialogues about strategic direction and prioritisation of investments.

⁴ Evaluative strategies encompass activities to support monitoring, evaluation and learning (MEL).

Introduction

The latest years' research focus on the evaluation of transformative innovation policy has been productive – yielding a growing number of published articles that provide new frameworks and ways of conceptualising transformative outcomes, as well as principles and proposed approaches to evaluation of transformation. The purpose of this paper is not to provide a comprehensive review of the literature⁵ (as has been done by e.g. Haddad et al. 2019), but rather to provide an overview and synthesis of some of the new knowledge and main lessons and proposals that policymakers can consider from the current academic state of the art.

A selection of 11 articles published in academic journals has been the main source for this synthesis (see Appendix A). Articles have been selected based on their focus on evaluation of transformative innovation policy or sustainability transitions⁶, their contribution in terms of proposing new frameworks or evaluation approaches⁷, and their timeliness (all articles are “latest developments” of researchers' work over time, and most articles have been published in the last three years). Earlier versions and related research have been used as a complement.

This synthesis is structured around four questions:

1. Who are the customers, and why do they want evaluations of transformation?

Who is demanding, and what are the main purposes and uses of evaluation of transformation? This drives an “evaluative strategy”, as well as the choice of approaches to monitoring, evaluation and learning (see #3).

2. What should get evaluated in transformative innovation policies?

What are transformative policies designed to do (what failures is the policy/instrument addressing)? What are expectations (set out in the program design/ToC)? What are characteristics/indicators or enablers of system transformation that are important to evaluate?

3. How should one approach TIP evaluation?

What are the different elements of an “evaluative strategy” for TIP? What principles and approaches are recommended at different points in a policy implementation cycle?

4. What is the unit of analysis being evaluated?

How to scope/delineate the boundaries of the system/long-term system transformation processes that are being evaluated? How to deal with multiple (nested) scopes and multi-level governance within a broader transformative process?

⁵ Stretching over studies of transformative and missions-oriented innovation policy, system innovation, sustainability transitions and socio-technical change, transitions management, evaluation, etc.

⁶ Keyword search (in abstracts) of: Evaluation and/or assessment and/or tracking of transformation and/or sustainability transitions and/or system change and/or system innovation

⁷ Seven of the articles propose frameworks and/or insights on “what” is evaluated in TIP (see Table 1), while the remaining four articles provide insights on evaluation approaches (“how”) and/or principles related to the evaluation of transformation.

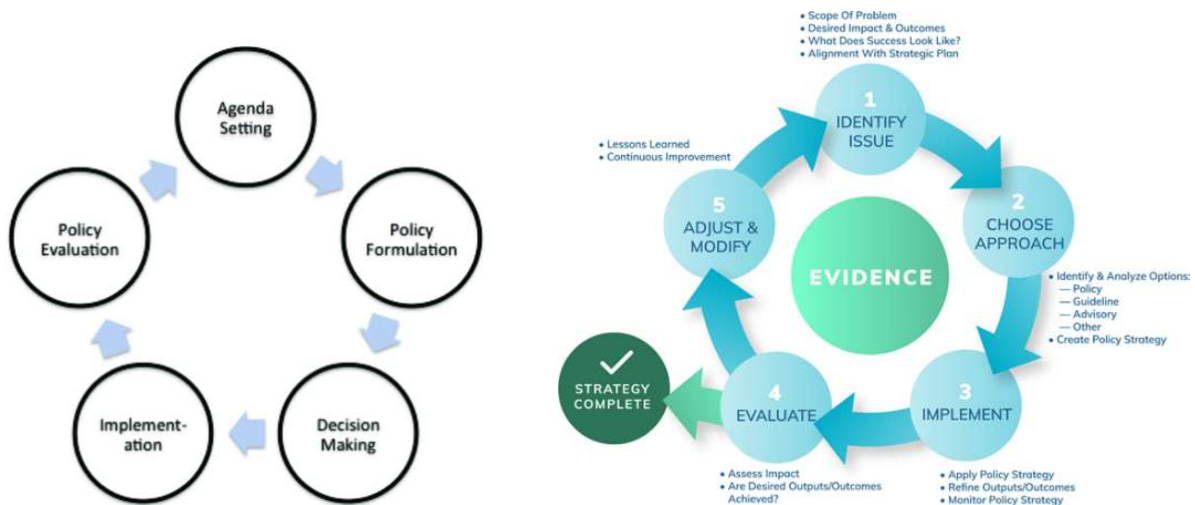
In the sections that follow, each question will be introduced, followed by a synthesis of the literature (with citations), and a short discussion of main lessons/takeaways for consideration in the design of an evaluative strategy for transformative innovation policy.

1. Who are the customers and why do they want to evaluate transformation?

In this document, we explore the evaluation of transformative innovation policy (or programmes) – where the main customers are understood to be organisations that invest public resources in a funding programme (or other sort of policy intervention) that aims to address societal challenges and generate broader impact (environmental, societal). Given the long-term, complex and interactive nature of these interventions, the customer is **not only the government or the agency through which public funds are channelled, but also the recipient initiative or platform that has been selected and given the mandate of leading the systemic change process**. Thus the customers of TIP evaluation are at various ‘levels of the investment chain’.

Evaluation in its broadest sense⁸ fulfils several purposes (providing accountability and ensuring public investments are yielding desired effects, developing understanding/knowledge of how and why things happen as they do, and providing input to next steps and improvements in the design and implementation of a policy action) and is an important part of the policy cycle. Design of an evaluative strategy (including monitoring, evaluation and learning activities) should be tailored to the policy in question, and should thus be an embedded part of policy design.

Figure 1: Example illustrations of the policy or policy strategy cycle



Sources: Howlett and Ramesh (1995) and www.chirobc.com

The different purposes of evaluation influence the approach and overall ‘philosophy’ for the evaluative strategy. There are two main approaches to evaluation: summative (focused on the outcome) and formative (focused on the process). A third approach, developmental evaluation (Patton 2010 and 2016), positions itself as “grounded in systems thinking, guiding adaptation to emergent and dynamic realities by collecting and analyzing real-time data to inform ongoing decision-making and

⁸ In this paper, we use the term “evaluation” in a broad sense to represent an evaluative strategy that includes monitoring, evaluation and learning activities (MEL).

implementation”. Whereas summative and formative evaluation approaches are primarily backward-looking and primarily targeted at programme funders and implementers, developmental evaluation is focused on informing the future and treats the intervention manager (a.k.a. process leader, transition manager or intermediary) as its primary customer. In essence, these three approaches can be mapped in relation to the three main purposes for evaluation.

Table 1: (Range of) Purpose, Focus and Approach to Evaluation

Purpose of evaluation	Providing accountability and ensuring public investments are yielding desired effects	Developing understanding/ knowledge of how and why things happen as they do	Providing input to next steps and improvements in the design and implementation of a policy action
Focus	Outcomes and impacts (in relation to goals)	Implementation processes	Real-time reflexive learning
Evaluation approach	Summative	Formative	Developmental

Evaluative strategies include elements of all three approaches, yet often there is more ‘weighting’ on a particular purpose and thus focus and approach to the evaluation. The literature highlights several main purposes for TIP evaluation, summarised in the table below.

Table 2: Purposes for TIP Evaluation

Purposes for TIP Evaluation	Citations from the literature
Providing evidence of contribution to systemic change in the desired direction	“results of an intervention can be no more than a contribution to (and not a determining cause of) the systemic changes being pursued” (Molas-Gallart et al. 2021)
Engaging participating actors in continual/real-time processes for learning and reflexivity, to ensure directionality (policy, programme and project levels) relate to relevant societal and ecological challenges	“collaborative practices characterizing TIPs are extended to evaluation and in so doing strengthen the collaborative feature in transition experiments and support organizational learning” (Molas-Gallart et al. 2021) “(evaluation scheme) serves as a basis for structured reflection and strategizing in support of experiments that help society to transition towards sustainability...to facilitate learning and accelerate progress across different experiments” (Luederitz et al. 2016) “an evaluation approach that provides input in the other policy cycle tasks...a reflexive layer to the policy definition and implementation process” (Molas-Gallart et al. 2021)
Leveraging learning from collective experimentation to guide iterative adjustments and development of the action(s) over time, to ensure continued relevance to directionality of change	“participants in experimental policy engagements (EPEs) need to learn whether the activity has set them up on the way to systemic transformation” (Molas-Gallart et al. 2021) “learning how to contribute to system innovation is the central goal of evaluation” (Molas-Gallart et al. 2021)

The main purposes for TIP evaluation are understanding and evidencing contributions to systemic change that have been made (summative), engaging participating actors in continual/real-time

processes for learning and reflexivity (formative), leveraging learning to guide iterative changes and ensure continued relevance to directionality for system change (developmental). These purposes point to a relatively heavier ‘weighting’ on formative and developmental (real-time reflection to inform forward-looking directions) approaches to evaluation. We return to this point in question three.

2. What should get evaluated in transformative innovation policies?

In the well-established 1st and 2nd generations of innovation policy (focused on knowledge production, products and processes, or on innovation systems), we have clear guidance on measures (or indicators) of research and innovation (in e.g. the Frascati and Oslo Manuals, OECD 2015 and OECD/Eurostat 2018) and how this data can be collected and used to gain insight and inform policy. The existing focus is on statistical measures of R&D inputs (in terms of funding sources, performers and personnel) and outputs (in terms of articles, patents and different kinds of innovation), complemented by assessments of how innovation systems operate and how changes in actor network composition/collaboration addressed system failures (captured through, for example, the Community Innovation Survey’s measures of innovation capabilities, the role of linkages with other firms and institutions in the innovation process, and external factors influencing innovation in firms, as well as measures for innovation activities and outcomes). The “what” that currently gets evaluated is primarily focused on the actors that produce knowledge and innovation (their capabilities, linkages and outputs, leading to longer-term impacts on economic performance, competitiveness and growth), with a relatively dimmer focus on the (enabling) innovation system. In the 3rd generation of innovation policy, the spotlight is more brightly focused on the system level – and how self-organising ecosystems of actors⁹ collaborate in order to achieve transformative change of a (socio-technical) system.

The focus is on a “network of actors who are willing to work on transformative change in an experimental setting...transformative change of socio-technical system...recognizing that all technical change has a directionality embedded in it, that actors need to become aware of this, and that there is a need to open up innovation spaces for more radical alternatives, involving more actors in this space and engaging them in a participatory way...then leads to carving out directionality portfolios and sustainable pathways for moving forward.” (Schot and Steinmueller 2018)

Increasing attention is placed on monitoring “transformative outcomes” in the form of changes in behaviour, relationships, activities or actions of people, groups, and organisations. This is a very different perspective (and set of indicators) on “what” should be evaluated. The academic literature¹⁰ presents a variety of conceptual frameworks and aspects that should be evaluated in transformative innovation policy. In this paper, these are grouped in four areas:

- system elements that are changing
- progression of system change
- directionality and strategic reorientation

⁹ With a relevant and representative composition of actor groups (corporates and startups, HEIs and RTOs, public sector, innovation support and civil society organisations), as well as a mandate/agency for change

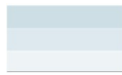
¹⁰ drawing from and combing approaches from technological innovation systems, sustainability transitions and multi-level perspectives in sustainability transitions, strategic niche, and transitions management, among others

- agency and transformative capacity

These aspects are included and treated in different ways in the various frameworks presented in the literature (see Table 3 below). A short description of each aspect follows.



System elements that are changing captures behavioural additionality on the system level (Haddad 2021) and encompasses the enabling conditions (processes or functions) that influence the ecosystem of actors' behaviours, as well as the "behaviour" (or strength) of the (innovation or socio-technical) system (at the start and how it develops over time). The technological innovation system (TIS) functions (Bergek et al. 2008, Hekkert et al. 2007)¹¹ are a well-established conceptualisation of these system elements and are used in Bergek and Haddad 2021, Bos et al. 2016 and Janssen 2019¹² frameworks. Frameworks for evaluating sustainability transition experiments (Luederitz et al. 2016; Williams and Robinson 2020) place focus on assessing the process – with a different conceptualization of aspects to assess (including inputs, enabling conditions, methods used and governance/leadership¹³).



Progression of system change captures signs of progression (or steps) along the change process at different phases of the evolutionary process of socio-technical change – from emergence of new niches, to expansion, to opening new regimes, as well as the contraction or destabilization of existing systems. "Signs" are indications of e.g. new capacities, structures, behaviours, norms, upscaling and uptake. "Progression" can be understood as how the signs build on each other (cumulatively) and contribute to new signs over time.

These steps are conceptualized as processes or interventions that influence changes in rules/institutions and guide changes in actors' behaviour, thus contributing to progression along a transformational journey. The phases of transformation are conceptualised based on multi-level perspectives (MLP) on transition (Geels 2002, Geels and Schot 2007, and many more). These steps have been coined "transformative outcomes" (Ghosh et al. 2021 and Bergek and Haddad 2021) or "markers of transition" (Williams and Robinson 2020), captured in "cumulative evaluations" (Luederitz et al. 2016).



Directionality and strategic reorientation captures both the initial strategic aims (or agenda) of the intervention (how it relates and contributes to confronting environmental and societal challenges) and how this is updated over time to reflect changes in the system and the broader context – i.e. reflexive learning on steps achieved so far and consideration of alternative paths forward from a diversity of perspectives (negotiating alignment and forming collective priorities).

Evaluation has traditionally related to intervention logics (or theories of change) – assessing the relevance, effectiveness and efficiency of the intervention in meeting its objectives (set out at the start). As transformative interventions are longer-term processes of systemic change (system innovation or

¹¹ Knowledge development, Knowledge exchange, Entrepreneurial experimentation, Market formation, Guidance of the search, Resource mobilization, Legitimation, Development of positive externalities

¹² The Janssen 2019 framework includes an approach to assess both policy design and to assess impact.

¹³ Which, in this paper, is included in "agency and transformative capacity"

transition pathways), the intervention logic needs to be continually revisited and ‘tweaked’ based on iterative processes of looking back (reflexivity) and looking forward (strategic foresight). Geels et al. (2016¹⁴) describe this process as “dynamics of strategic reorientation”. Robinson et al. (2021) discuss the importance of bridging the use of foresight (as a systematic process of developing strategic intelligence, mobilising understandings of the past and structured articulations of potential futures) to aid decision making. Evaluations need to capture this dynamic perspective of directionality. In the articles included in this synthesis, these aspects are conceptualised as “flexible theories of change” (Molas-Gallart et al. 2021), use of “directionality filters” to interrogate the functions (Bergek and Haddad 2021), or use of actor-based/participatory processes to consider external/landscape factors and translate strategic ambitions into concrete actions going forward (Bos et al. 2016 and Kroll 2019).



Agency and transformative capacity captures the institutional set-up, strength and functioning of the “driving force forward” which has been conceptualised in terms such as “systemic intermediary” (van Lente et al. 2003), “systemic intermediation” (Kanda et al. 2020), “transition intermediary” (Hysalo et al. 2018) or more generally “intermediaries in sustainability transitions” or transformation (Kivimaa et al. 2019a and 2019b). Intermediaries provide “transformative and orchestrating capacities” (Hölscher et al. 2018) – coordinating multi-actor processes across scales and sectors to facilitate and align experimentation and collective visioning, as well as catalysing transition activities by qualifying and aggregating market information, articulating demand and operationally facilitating experimentation.

Although the seminal articles introducing TIP highlight “the need for anticipation, experimentation, learning, and the formation of bridging networks and alliances suggest(ing) new institutional arrangements and governance structures that cut across governments, markets, and civil society” (Schot and Steinmueller 2018, p.1564), this aspect of agency/governance is not strongly highlighted in the suggested frameworks for TIP evaluation. Turnheim et al. (2015) explain that “integrated appraisal approaches should seek to contribute to policy and governance of sustainability transitions” and compare how three different analytical approaches (quantitative systems modelling, socio-technical transition analysis, and initiative-based learning) address the aspect of governing transitions.

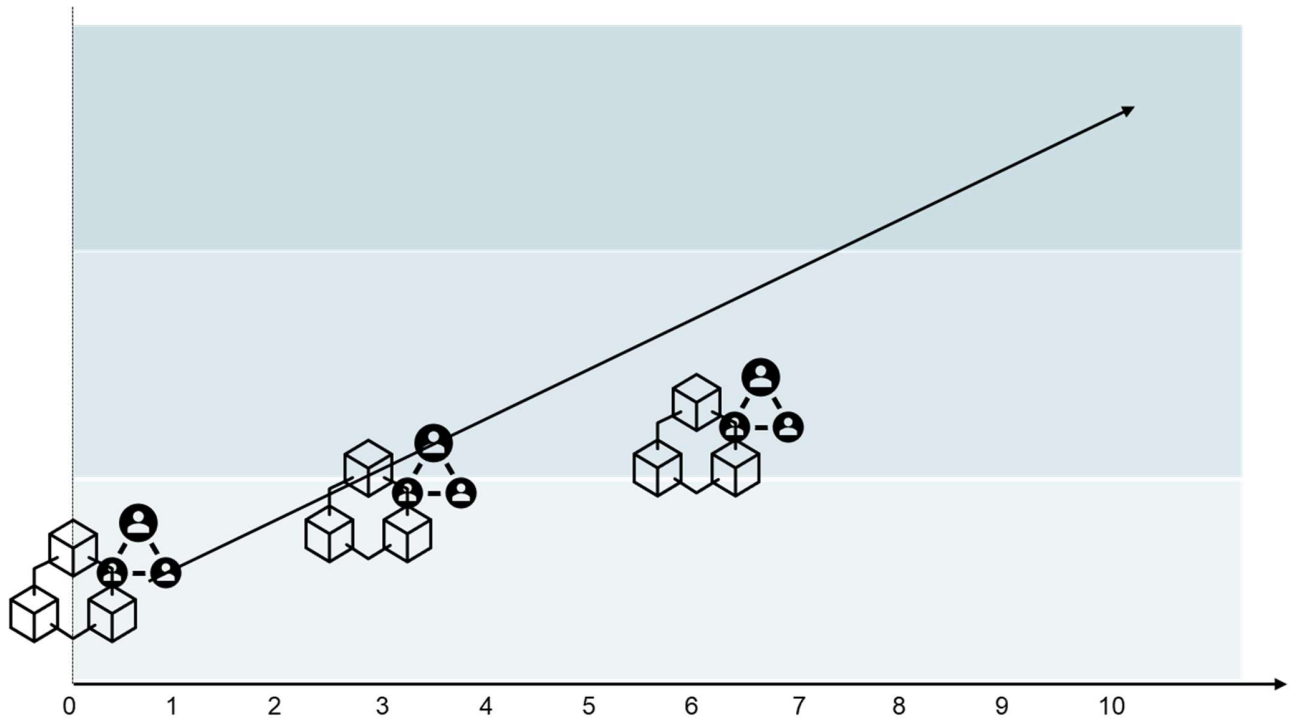
Frameworks for evaluation of sustainability transition experiments (Luederitz et al. 2016 and Williams and Robinson 2020) include assessments of governance, facilitation methods, etc. as part of process evaluation. And Janssen 2019 highlights “the importance of visible, responsible and competent leadership, transparency and a pro-active communication strategy, mechanisms for learning and changing course” as part of evaluation of policy design. Kanda et al. (2019) describe the role of intermediaries in relation to TIS functions, and Kivimaa et al. (2019a) present a classification of innovation intermediaries’ functions and activities, differentiated by level (niche, niche-regime, regime), type of intermediary and phase of transition (pre-development & exploration, acceleration & embedding, stabilisation, destabilisation). These (and other recent academic contributions) could inform an assessment of this aspect of TIP.

The combination of the various “what’s” that should be considered and assessed in an evaluative strategy is illustrated in the Figure 2 below. This includes an assessment of the ‘strength’ and functioning

¹⁴ Based on Geels (2014)

of the innovation and socio-technical system, the direction and reorientation of the transformative change that is intended (representing the path towards a particular goal/mission or contribution to a broader system-level transformation¹⁵), the capacity for driving transformative change processes forward¹⁶, and development steps or progression of change that is achieved over time.

Figure 2: Illustration of “what” should be assessed in TIP evaluative strategies



It should be noted that the relative strength of the system/systemic intervention and its transformative capacity, the clarity (and ambition) of its strategic direction, and its position relative to phases of socio-technical change processes at “time 0” can vary. All systemic interventions are influenced by their context and usually have a history. Depending on the policy/programme in question, expectations at “time 0” will vary.

Existing evaluative strategies for innovation policy/programmes include various aspects of this set of “what’s” (e.g. assessment of collaborative groupings/system resources, strategic intentions, innovation leadership and implementation capacity, and outputs/outcomes/impacts over various phases of time). The key differences in TIP evaluation are a focus on the system/system-level innovation (vs. a focus on actors), a directionality that addresses system-level societal challenges (vs. innovation or innovative capacity more broadly), and signs of progress (outcomes) that evidence the contribution of the collective efforts towards the broader transformative aims (vs. a causal link to a particular solution or broader innovative capacity).

¹⁵ determined by the policy need and ToC set at the start of the programme/project/EPE

¹⁶ leadership/mgmt./governance, including responsibility for processes like strategic intelligence, reflection, etc.

Table 3: Overview of aspects to be evaluated in transformative innovation policy – excerpts from theoretical frameworks

	Bergek and Haddad 2021	Bos et al. 2016	Ghosh et al. 2021	Janssen 2019	Kroll 2019	Luederitz et al. 2016	Williams and Robinson 2020
System elements (processes or functions)	<ul style="list-style-type: none"> - Knowledge development and diffusion - Entrepreneurial experimentation - Market formation - Guidance of the search - Resource mobilization - Legitimation - Development of positive externalities 	<p>TIS approach useful for study of the conditions for success or failure of innovation trajectories</p> <ul style="list-style-type: none"> - Entrepreneurial activities - Knowledge development - Knowledge diffusion through networks - Guidance of the search - Market formation - Resource mobilization - Creation of legitimacy 		<ul style="list-style-type: none"> - Guiding the direction of search - Knowledge development - Knowledge exchange - Entrepreneurial experimentation - Resource mobilization - Legitimation - Market creation 		<p><i>Not directly addressed in this model.</i> Article focuses on evaluation of sustainability transition experiments, and uses traditional logic mode scheme to assess the experimental process:</p> <ul style="list-style-type: none"> - Inputs invested into experiments - Processes performed in experiments - Outputs generated by experiments (e.g. built capacities, actionable knowledge, accountability, structural changes) - Outcomes accomplished by experiments (e.g. socio-ecological integrity, resource maintenance, etc.) 	<p><i>Not directly addressed in this model.</i> Article proposes a framework (and indicators) for evaluation of sustainability transition experiments (STEs), assessing three levels:</p> <ol style="list-style-type: none"> 1) Process – fairness and inclusivity of the process, the quality and appropriateness of tools and methodologies used, and the adaptive and reflexive capacity 2) Societal effects – short-term outputs and medium-term outcomes of an STE (e.g. individual capacity, usable products, networks and relationships, institutional change) 3) Sustainability transition impacts of STEs (organized into five characteristics of development pathways)
Progression/ degree of system change (transformative outcomes in different stages)	<p>Transformative processes related to actor networks at niche- and regime-levels</p> <ul style="list-style-type: none"> - Entry of new actors 	<p>Likelihood of “actors fulfilling relevant function” by mapping perceptions of innovation in terms of type (incremental, modular, architectural, radical)</p>	<p>TOs are (12) processes or interventions that lead to deeper changes in sets of rules that guide actors, grouped in</p>	<p>Assessing progression of the system by interrogating each function with:</p> <ul style="list-style-type: none"> - How important was function j for 		<p>Sustainability transition experiments have an objective to initiate and facilitate radical long-term transitions through orchestration of</p>	<p><i>Not directly addressed in this model.</i> Assessment of societal effects captures the link between the design and execution of</p>

<p>of niche emergence, expansion and opening up new regimes)</p>	<ul style="list-style-type: none"> - Formation of new knowledge/ technology/ business networks - Configuration and de-configuration of political networks - Development of political capacity and change advocacy <p>Transformative processes related to institutions on niche- and regime-levels</p> <ul style="list-style-type: none"> - Articulation of visions and expectations - Framing and redefinition of values, norms, and practices - Mobilization and de-mobilization of political support - Introduction of new regimes 	<p>and value at different levels of MLF</p> <p>Multi-level framework helps to understand transition dynamics on three levels of analysis:</p> <ul style="list-style-type: none"> - Micro level with niches (TIS functions 1-3) - Meso level with sociotechnical regimes (TIS functions 3-6/7) - Macro level with elements of the landscape <p>Also consider actor specifics in terms of role (e.g. financial, research, producer, supplier, user, etc.) and dimension in system (technology, user practices, application domains, policy, etc.)</p>	<p>three macro-processes:</p> <p>1)Building and nurturing niches</p> <ul style="list-style-type: none"> - Shielding - Learning - Networking - Navigating expectations <p>2)Expanding and mainstreaming niches</p> <ul style="list-style-type: none"> - Upscaling - Replicating - Circulating - Institutionalising <p>3)Opening up and unlocking regimes</p> <ul style="list-style-type: none"> - De-aligning and destabilising - Unlearning and deep learning in regimes - Strengthening regime-niche interactions - Changing perceptions of landscape pressures 	<p>the success of the TIS?</p> <ul style="list-style-type: none"> - What did the government do to strengthen function j? How substantial was this input? - How effective was policy in strengthening function j? What output was achieved? - How decisive was policy for strengthening function j? - Did solving function j strengthen other functions? 		<p>specific experiments that aim to challenge the status quo.</p> <p>Evaluate how outputs from specific experiments (capacities, knowledge, accountability, structural changes, uptake of experiments) contribute to outcomes in terms of sustainability related accomplishments</p> <p>Evaluation of experiments that aim at larger goals (sustainability transition) requires cumulative evaluations</p>	<p>process(es) and outcomes.</p> <p>Discuss (in conclusions) the concept of ‘markers of transition’ – i.e. using process characteristics and societal effects to anticipate and assess if certain types of impacts (markers of transition) are present.</p>
<p>Directionality and strategic reorientation</p>	<p>Each of the three sets of transition-related processes (sociotechnical system functions, processes related to changes in actor networks and in institutions) should be scrutinized from a directionality point of view to determine whether they contribute strengthening the existing sociotechnical</p>	<p>Not directly addressed in this model. However, strategic reorientation explored through consideration of landscape factors and their possible influence on changes to the regime.</p> <p>Possible landscape factors mentioned in the article: demographic factors, macroeconomic</p>	<p>From ‘sister article’ Molas-Gallart et al. 2021: ...recommend using a “flexible theory of change” that is revisited and redefined as a result of the formative evaluation process</p> <p>The ToC will be used to foster (first and second order) learning and reflexivity among participants and to</p>		<p>Article proposes a framework for evaluating innovation strategies with a transformative ambition and focuses on reflexive and actor-based processes to translate strategic ambitions into concrete actions and maintaining coherence between these three levels:</p>	<p>Ultimate aim of experiments is sustainability-related outputs that contribute to radical transitions. Actions in experimentation include:</p> <ul style="list-style-type: none"> - Defining a baseline and goal for the intervention - Creating a specific set-up to administer interventions 	<p>Use the SDGs to assess the direction and comprehensiveness of development path change</p>

	configuration, the development of new configurations, or both.	factors, environmental factors, geographical factors, societal factors.	help assess if the policy is contributing to move towards its objectives.		<ul style="list-style-type: none"> - Strategic agenda setting (on the basis of shared frames and narratives) - Thematic orientation (and constitution of directionality, anchored) - Actual implementation and instrumentation 	<ul style="list-style-type: none"> - Measuring the effects of the interventions against the baseline - Evaluating the effects against sustainability criteria - Offering evidence-supported recommendations on how to implement results 	
Agency and transformative capacity				As part of assessing the policy design (vs. impact) evaluation, Janssen explains the importance of visible, responsible and competent leadership, transparency and a pro-active communication strategy, mechanisms for learning and changing (policy) course, as well as high level support and capable implementing agencies		Not directly addressed in this model. However, a part of the process for experimentation is "creating a set-up to administer interventions"	<p>Among the process evaluation categories to be assessed is governance (including stakeholder capacities, power relations, engaging future and non-human actors, recognition) and methods (including dialogue, negotiation, collective problem solving, reflexivity)</p> <p>Among the impact categories is socio-technical systems and governance exploring governance role and relationships (including inclusion of new actors and issues, changes in decision-making or who gets to participate in decision making that affects the STE</p>

3. How should one approach TIP evaluation?

While the literature highlights many challenges to evaluating systemic transformation (e.g. complex and long-term processes make it difficult to assess additionality of a temporal and spatially delimited intervention), the literature also proposes a number of guiding principles and proposed approaches for TIP evaluation.¹⁷ Some commonalities and key messages across the literature (as well as citations and additional references) are summarised in the table below. Additional detail and excerpts from articles is presented in Appendix II.

Table 4: Guiding principles and proposed approaches for TIP evaluation

Principles and approaches for TIP evaluation	References and citations from the literature
Integrate evaluation (an evaluative strategy) as an integrated part of policy design	“...Evaluation emerges as a core activity in transitions, periodically informing experiments to adapt, extend and revise the envisioned pathway. To achieve this requires: ex-ante evaluation prior to the implementation of experiments to inform their design; formative evaluation to adjust and improve ongoing experiments; and, ex-post evaluation to appraise the contribution of experiments to sustainability after completion.” (Luederitz et al. 2016)
Adopt formative, developmental and realist evaluation practices, including a flexible theory of change, real-time observations, and participatory techniques for iterative and reflexive process monitoring	<p>Formative/constructive monitoring (Molas-Gallart et al. 2021)</p> <p>Interactive Learning and Action evaluation approach (ILA) as developed by Hoes et al. (2010), and a corollary monitoring approach, namely ‘reflexive process monitoring’ developed by Van Mierlo et al. (2010)</p> <p>Action-oriented transformation research...third loop learning (Fazey et al. 2018)</p> <p>Client-oriented evaluation of policy mix using qualitative methods to analyse supporting functions and destabilising policies to understand perceptions of policy results (Kivimaa et al. 2017)</p>
Ensure iterative and reflexive monitoring is used to support learning and ongoing change processes (on project, programme and policy levels)	<p>“Iterative and reflexive monitoring and evaluation needs to be an integral part of sustainability transition experiments to support individual and organizational learning promoting ongoing change and up-scaling impact.” (Luederitz et al. 2016)</p> <p>Evaluations scrutinize assumptions, structures, and values as well as related and unrelated changes in society in order to inform future actions (Schot and Geels, 2008; Loorbach and Rotmans, 2009). Embedded within these different modes of evaluation are reflexive learning processes which continually assess the transformational potential of experiments and the evaluation itself.</p>
Combine a mix of methods and analytical approaches to gather intelligence and track change over different periods of time	“A system undergoing transformation is never fully ordered and stabilised, and so requires a continuing readjustment in the categories and metrics used to describe and analyse it.” Evaluation needs to provide: understanding of recent past and present, lessons about scaling experimentation, and ability for developing future projections. (Turnheim et al., 2015)

¹⁷ Haddad 2021 recommends a set of 7 steps for TIP evaluation.

Use a nested approach to assess multiple levels of transformative initiatives	Evaluate how outputs from specific experiments (capacities, knowledge, accountability, structural changes, uptake of experiments) contribute to outcomes in terms of sustainability related accomplishments/larger goals...requires cumulative evaluations (Luederitz et al. 2016)
---	---

There is an apparent consensus around the overall philosophy and framing of TIP evaluation¹⁸ (including the importance of understanding the context within which an intervention takes place and relating this to a flexible/iteratively developed theory of change), as well as the ‘weighting of’ a formative and developmental approach (including the importance of processes for real-time, iterative and participatory reflection and learning). Furthermore, there is apparent consensus on the need for integrating an evaluative strategy (i.e. processes for monitoring, evaluation and learning) in the programme design, and using a mix of methods to assess transformation across different levels of intervention and periods of time. This relates to the ‘unit of analysis’ being evaluated, to which we turn in the next section.

4. What is the unit of analysis being evaluated? What scope of evaluation?

In evaluation of TIP, it is important to delineate the system and scope of system change that is being evaluated (in terms of e.g. policy/governance levels, sectors/themes, time/duration and spatial characteristics of the change process). The literature included in this synthesis presents perspectives on several different levels and scopes of evaluation.

In terms of policy levels, Kroll (2019) focuses on policy strategy. Molas-Gallart et al. (2021) distinguish between the different levels of policy action and propose establishing a connection between policy or programme level (higher level) and lower-level interventions. And Ghosh et al. (2021) refer to these (lower-level) interventions as “Experimental Policy Engagements”. EPEs are “diverse ways in which STI policymakers engage with processes of societal experimentation for sustainable transformation: initiating, supporting or mobilising, and evaluating such initiatives for informing decision-making, enabling processes of social learning, developing alternative pathways and enacting desirable futures”.

Molas-Gallart et al. (2021) explain that EPEs aim at making unfolding transformation processes more transformative and become experimental because they are time-bounded attempts to influence the transformation in a reflexive and learning-oriented manner. This conceptualisation of scope is similarly adopted in sustainability transition experiments (Luederitz et al. 2016; Williams and Robinson 2020), where sustainability transition initiatives are considered as small-scale networks, creating an interface between actors from various institutional realms, and in doing so fostering learning through participation and interaction.

Molas-Gallart et al. (2021) further explain that taking a systemic view requires a shift in interest from the project or programme levels to the level of the whole system that the policy initiatives are trying to effect (Caffrey and Munro 2017). **The evaluators’ main focus may thus change from the analysis of a specific intervention, to the study of the effects of portfolios of interventions and the systemic impacts of policy mixes involving an integrated evaluation of the different policy instruments and their**

¹⁸ Within “realist evaluation” (Pawson and Tilley 1997) – where effectiveness is not determined by outcomes alone (typical cause-effect evaluation). Rather, there is a consideration of the theoretical mechanisms applied (outlined in a theory of change) and the socio-historical context in which the policy action/experiment is implemented. Thus, evaluations consider context-mechanism-outcome (C-M-O).

interactions (Magro and Wilson 2013) – i.e. using a nested approach to assess multi-level TIPs, as well as broadening the scope of actors involved (beyond the R&I system) to include government and a wide range of other stakeholders (Arnold and Barker forthcoming 2022).

In addition to needing a nested approach to evaluation that encompasses “lower-level” experiments in relation to “higher-level” programme and policy levels, evaluation of TIP needs to address the time and spatial characteristics of system-level change processes as they unfold. Turnheim et al. (2015) explain that socio-technical transitions involve change unfolding over extended periods of time (decades) and spanning different scales (territorial, jurisdictional, organisational, cultural) and compare how three different analytical approaches (quantitative systems modelling, socio-technical transition analysis, and initiative-based learning) address the aspect of scale (ranging from national and sector-scale to local scale) and temporality (ranging from multiple decades to 5-15 years).

Key differences between the three approaches.

		Quantitative systems modelling	Socio-technical analysis	Initiative-based learning
Scale and temporality	Analytical scale	Various – but often global, national and sector-scale analysis	Mainly national systems, with comparisons across countries	Local scale, sometimes with comparisons across contexts
	Multi-scale linkages	Linkages between sectoral models at different scales	Landscape-regime-niche links processes at different scales and temporalities	Not explicit
	Time horizon	Long-term perspective (decades)	Long-term perspective (decades)	Short timeframe (5–15 years)
	Time orientation Temporal articulation	Future scenario projections Current decisions informed by future projections	Historical and ongoing transitions Cases approached via pattern recognition over time	Transitions in the making Decision-making and implementation temporality

Source: Turnheim et al. (2015: 245)

Turnheim et al. (2015) go on to highlight that “sustainability transitions can only be steered in real time, over time, with knowledge of past developments. **Steering transitions requires a particular kind of sense making: the ability to “zoom in and out” between levels of analysis and to ‘zip back and forth in time’.** This can be done by linking the different perspectives and the multiplicities of scale and temporalities they offer.” Molas-Gallart et al. (2021) similarly highlight that “even without engaging in an extensive systemic analysis, assessing the results of policy interventions remains a very challenging task because the changes sought will typically occur a long way downstream from the intervention and be the result of a complex interaction of factors which may or may not be directly related to the interventions (ibid., p.5). Leuderitz et al. (2016) also acknowledge the challenge of evaluating “comprehensively” when the evaluation is focused on particular interventions, projects or experiments (even if the experiments aim at a larger sustainability transition goal) and recommend **cumulative evaluations** to capture how outcomes from experiments complement and reinforce outcome features in other/parallel experiments (targeting the same longer-term aims).

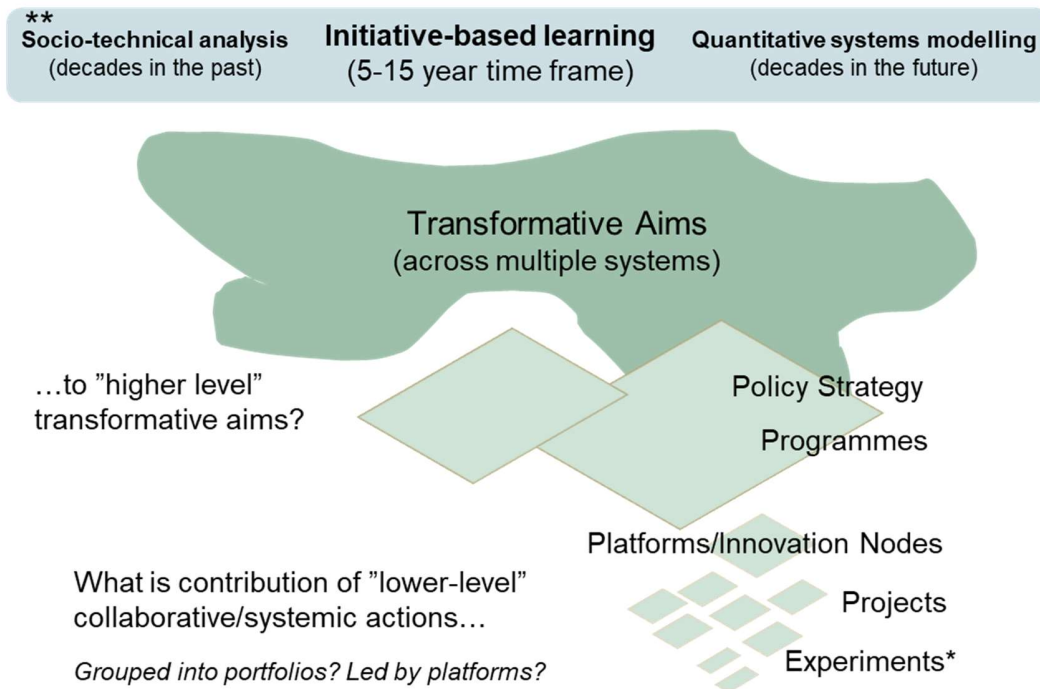
The literature reviewed in this synthesis highlights the challenge of scoping an evaluation (or developing an evaluative strategy) that encompasses:

- the various (nested) levels of system change and governance (from the very concrete projects and experiments, to the meso-level programmes and policy areas, and the macro “total system performance”),
- capturing and following outcomes (and downstream activities or “ripple effects” of these over time), and
- describing how outcomes in one experiment or systemic initiative relate to efforts going on in other systemic initiatives

The literature reviewed also has apparent consensus around a focus on a short-to-medium term time frame (5-15 years) for **initiative-based learning** (see previous sections describing analytical approaches), with the use of other analytical approaches to assess other time frames (historically or future scenarios) for system change processes.

This “scoping” of TIP evaluation is illustrated in Figure 3 below.

Figure 3: Scoping of TIP evaluative strategies



Conclusions and Next Steps

This summary overview of the current academic “state of the art” on the evaluation of transformative innovation policy has provided some insights and guidance on the customers and main purposes of these evaluations, the recommend approach to TIP evaluation, the different aspects (the “what”) of system change that should be evaluated, and the unit of analysis/scope of TIP evaluation.

The main purposes for TIP evaluation are understanding (and evidencing) contributions to systemic change that have been made; engaging actors in continual/real-time processes for learning and reflexivity; and leveraging learning to guide iterative changes (and ensure continued relevance) to directionality for system change. The literature presents a relative consensus around the overall philosophy and framing of TIP evaluation (in realist evaluation), as well as the ‘weighting of’ a formative and developmental approach (including the importance of processes for real-time, iterative, and participatory reflection and learning to inform forward-looking directions). Furthermore, there is apparent consensus on the need for integrating an evaluative strategy (i.e. processes for monitoring, evaluation and learning) in the programme design, and using a mix of methods to assess transformation across different levels of intervention and periods of time.

Although presented in a variety of conceptual framings, the academic literature highlights four key aspects of system change/transformation that should be assessed in TIP evaluation:

- the goal-setting (directionality and strategic orientation),
- the system elements that are changing,
- the progression (depth/scope) of system change, and
- the momentum provided through agency and transformative capacity.

The literature recommends carefully choosing and planning for the evaluative strategy from the start, using cumulative and nested evaluations to track change processes over time and across initiatives/ experiments on different levels, and linking different perspectives and multiplicities of scale and time through “particular kinds of sense-making” processes.

So what does this mean for policymakers? What are some of the key implications and considerations that can be gleaned from the academic theory?

It is important to **design “with the end in mind”** – i.e. design transformative (or mission-oriented) programmes with an evaluative strategy in mind from the start.

With **relative consensus in the literature** around the questions of “For whom and why”, “How” and “What unit of analysis”, this **provides some initial “helpful inputs” to planning an evaluative strategy**, which should include:

- approaches for capturing, participatory reflection and learning from ongoing system change processes on micro level;
- approaches for broader ‘zooming out’ (across portfolios of experiments/initiatives at different levels and scales) to assess the cumulative contributions and progress towards longer-term system transformation aims; and
- approaches to inform forward-looking directions (e.g. prioritising among alternative paths of action within an experiment/initiative, or prioritising among alternative experiments)

There are multiple ways of framing and describing (and thus evaluating) the “what” of system transformation – **no common framing or commonly accepted “signs of transformation” to look for and assess**. In fact, existing literature recommends different aspects or perspectives to examine¹⁹ as well as different analytical approaches (Turnheim et al. 2015) to assessing socio-technical systems change with different time perspectives. This may imply one lens and approach for assessing the system change process in real-time (i.e. monitoring processes) and another lens and approach for assessing the system change process after it has occurred (i.e. interim and ex-post evaluation).²⁰ The development of a

¹⁹ Whereas the literature in this synthesis focuses on technological innovation systems (TIS) functions and processes, other literature (Edler et al. 2021; Ghosh and Schot 2019; Mickwitz et al. 2021) recommend other dimensions for evaluating system transformation (e.g. Science and technology, Policy and government, Market and users, Industry structure and strategy, Socio-cultural).

²⁰ Janssen, Bergek and Wesseling (2022) propose such a “delineation in lenses”, recommending the combination of summative evaluations with more frequent formative evaluations.

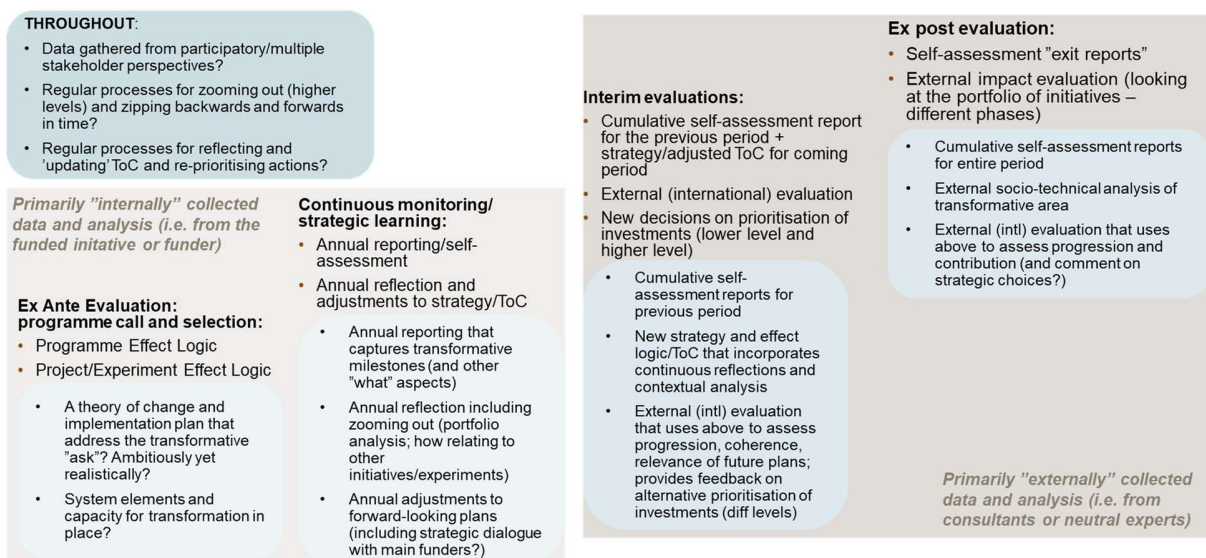
common way of understanding, describing and assessing transformative change processes that are supported by various policies within and across various funding agencies may be desirable.

Given that transformative innovation policies (3rd generation) add on to (and do not replace) existing research and innovation policies (in 1st and 2nd generations), the design of evaluative strategies for TIP implies additional challenges. **How can one realistically “add on” new evaluation needs without over-burdening the transformative interventions that are being evaluated? This will likely require new approaches** (including e.g.: adopting standard monitoring, evaluation and learning ‘protocols’ across funding agencies, leveraging user-friendly templates and digital interfaces, connecting to existing data/databases in order to “auto fill” and coalesce/cumulate information).

The importance of longer-term, cross-portfolio and policy area perspectives, coupled with the need to continually iterate on strategic direction and prioritisation of investments (towards longer-term transformative aims) **implies a different (more active) role for funding and policymaking agencies** and a potentially more prevalent engagement of the political level of government (where bigger decisions and prioritisations among investments and challenges may need to be anchored – see e.g. Arnold and Barker forthcoming 2022).

The literature in this synthesis provides several over-arching insights and guidance on TIP evaluation. Other existing literature provides recommendations on system-level evaluations (Arnold 2004; Edler et al. 2008) and evaluations of policy mixes across multi-level and multi-actor contexts (Flanagan et al. 2011; Magro and Wilson 2013). Although providing helpful insights and guidance, the literature does not (yet) provide concrete (empirical) examples of how TIP evaluation can be done in practice. What does an evaluative strategy for TIP look like? What are the different elements and approaches needed for monitoring, evaluation and reflective learning? A preliminary sketch of elements and approaches to include in an evaluative strategy for TIP is presented in Figure 4 below.

Figure 4: Elements and approaches in an evaluative strategy for TIP



After establishing a strategy for monitoring, evaluation and learning, additional questions emerge. How can an evaluative strategy be implemented in practice? Who should be asking which questions from whom and when? Some clues can be found in evaluating climate change. And some clues and lessons can be extracted from existing monitoring, evaluation and learning practices. This (i.e. learning from existing practice) is the next step of GReaTr exploration – working towards a theoretically-informed and practically-feasible evaluative strategy for TIP.

Appendix I: List of articles

- Bergek, A. and Haddad, C.R. (2021). 'Evaluating transformative innovation policy outcomes as unfolding processes of change in sociotechnical configurations' Book chapter manuscript (Transformative Metrics) currently under peer review
- 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 (4): 185–214.
- Ghosh, B., Kivimaa, P., Ramirez, M., Schot, J. and Torrens, J. (2021). 'Transformative outcomes: assessing and reorienting experimentation with transformative innovation policy', *Science and Public Policy*, 2021/00: 1-18. DOI: 10.1093/scipol/scab045.
- 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.
- Kivimaa, P., Hyysalo, S., Boon, W., Klerkx, L., Martiskainen, M., and Schot, J. (2019a). 'Passing the baton: How intermediaries advance sustainability transitions in different phases', *Environmental Innovation and Societal Transitions*, 31: 110-125.
- Kroll, H. (2019). 'How to evaluate strategies with a transformative ambition? A proposal for a structured, process-based approach', *Science and Public Policy*, 46 (5): 635-647. DOI: 10.1093/scipol/scz016.
- Luederitz, C. et al. (2016). 'Learning through evaluation – A tentative evaluative scheme for sustainability transition experiments', *Journal of Cleaner Production*, (2016): 1-16.
- Molas-Gallart, J., Boni, A., Giachi, S. and Schot, J. (2021). 'A Formative Approach to the Evaluation of Transformative Innovation Policies', *Research Evaluation*, 2021: 1-12. DOI: 10.1093/reseval/rvab016.
- Patton, M.Q. (2021). 'Evaluation Criteria for Evaluation Transformation: Implications for the Coronavirus Pandemic and the Global Climate Emergency', *American Journal of Evaluation*, 42 (1): 53-89.
- Turnheim, B. et al. (2015). 'Evaluating sustainability transitions pathways: Bridging analytical approaches to address governance challenges', *Global Environmental Change*, 35 (2015): 239-253.
- Williams, S. and Robinson, J. (2020). 'Measuring Sustainability: An evaluation framework for sustainability transition experiments', *Environmental Science and Policy*, 103 (2020): 58-66.

Additional References

- Arnold, E. (2004). 'Evaluating research and innovation policy: a systems world needs systems evaluations', *Research Evaluation*, 13 (1): 3-17.
- Arnold, E., Åström, T., Glass, C. and De Scalzi, M. (2018). *How should we evaluate large, complex programmes to support innovation and socio-technical transitions?* Technopolis Group.
- Arnold, E. and Barker, K. E. (2022). What past changes in Swedish policy tell us about developing third-generation research and innovation governance. In: Benner, M., Schwaag-Serger, S. and Marklund, G. (Eds.) *Towards a smart society? Innovation policy and the challenges of social inclusion, environmental resilience and sustainable growth*. Cheltenham: Edward Elgar.
- Åström, T. and Arnold, E. (2020). *Metautvärdering av första omgången strategiska innovationsprogram efter sex år – Appendix A*. Vinnova Rapport VR2020:10, Stockholm: Vinnova.

- Åström, T., Arnold, E. and Olsson, J. (2020). *Metautvärdering av andra omgången strategiska innovationsprogram efter sex år – Appendix A*. Vinnova Rapport VR2020:19, Stockholm: Vinnova.
- Åström, T., Arnold, E. and Olsson, J. (2021). *Metautvärdering av tredje omgången strategiska innovationsprogram efter sex år – Appendix A*. Vinnova Rapport VR2021:10, Stockholm: Vinnova.
- Bergek, A., Jakobsson, S., Carlsson, B., Lindmark, S. and Rickne, A. (2008). 'Analysing the functional dynamics of technical innovation systems: A scheme of analysis', *Research Policy*, 37 (3): 407-429.
- Boni, A., Giachi, S. and Molas-Gallart, J. (2019). *Towards a Framework for Transformative Innovation Policy Evaluation*. Transformative Innovation Policy Consortium (TIPC) Research Report, April 2019.
- Caffrey, L. and Munro, E. (2017). 'A systems approach to policy evaluation', *Evaluation*, 23 (4): 463-478.
- Chataway, J., Daniels, C., Kanger, L., Schot, J., and Steinmueller, E. (2017). 'Developing and Enacting Transformative Innovation Policy'. Paper presented at the 8th *International Sustainability Transitions Conference*, Gothenburg, Sweden, 18 – 21 June.
- Edler, J., Ebersberger, B. and Lo, V. (2008). 'Improving policy understanding by means of secondary analyses of policy evaluation', *Research Evaluation*, 17 (3): 175-186.
- Edler, J., Köhler, J., Wydra, S., Salas-Gironés, E., Schiller, K., Braun, A. (2021). 'Dimensions of systems and transformations: Towards an integrated framework for system transformations', Fraunhofer ISI Working Paper Sustainability and Innovation No. S 03/2021.
- Fazey, I, Schöpke, N, Caniglia, G et al. (45 more authors) (2018). 'Ten essentials for action-oriented and second order energy transitions, transformations and climate change research', *Energy Research and Social Science*, 40: 54-70.
- Flanagan, K., Uyarra, E. and Laranja, M. (2011). 'Reconceptualising the 'policy mix' for innovation', *Research Policy*, 40: 702-713.
- Geels, F.W. (2002). 'Technological Transitions as Evolutionary Reconfiguration Processes: A Multi-Level Perspective and a Case-Study', *Research Policy*, 31: 1257–74.
- Geels, F.W. (2014). 'Reconceptualising the co-evolution of firms-in-industries and their environments: developing an inter-disciplinary triple embeddedness framework', *Research Policy*, 43 (2): 261–277.
- Geels, F.W., Kern, F., Fuchs, G., Hinderer, N., Kungl, G., Mylan, J., Neukirch, M. and Wassermann, S. (2016). 'The enactment of socio-technical transition pathways: A reformulated typology and a comparative multi-level analysis of the German and UK low-carbon electricity transitions (1990-2014)', *Research Policy*, 45(2016): 896-913.
- Geels, F. and Schot, J. (2007). 'Typology of sociotechnical transition pathways', *Research Policy*, 36: 399-417.
- Ghosh, B. and Schot, J. (2019). 'Towards a novel regime change framework: Studying mobility transitions in public transport regimes in an Indian megacity', *Energy Research & Social Science*, 51: 82-95.
- Haddad, C.R. (2021). *Evaluating transformative innovation policy: towards an integrated framework* [Licentiate thesis]. Department of Technology Management and Economics, Chalmers University of Technology.

- Haddad, C.R., Nakić, V., Bergek, A. and Hellsmark, H. (2019). 'The policymaking process of transformative innovation policy: a systematic review'. Paper presented at 4th *International Conference on Public Policy (ICPP4)*, Montréal, Canada, 26 – 28 June.
- Hekkert, M.P., Suurs, R.A.A., Negro, S.O., Kuhlmann, S., and Smits, R.E.H.M. (2007). 'Functions of innovation systems: A new approach for analysing technological change', *Technological Forecasting and Social Change*, 74 (4): 413-432.
- Hoes, A-C., Regeer, B. and Bunders, J. (2010). *Facilitating Learning in Innovative Projects: Reflections on our experiences with ILA-monitoring*.
- Hölscher, K., Frantzeskaki, N. and Loorbach, D. (2018). 'Developing Transformative and Orchestrating Capacities for Climate Governance Experimentation in Rotterdam' in Turnheim, B., Kivimaa, P. and Berkhout, F. (Eds) *Innovation Climate Governance: Moving Beyond Experiments*, Cambridge University Press.
- Howlett, M. and Ramesh, M. (1995). *Studying Public Policy: Policy Cycles and Policy Subsystems*. Oxford: Oxford University Press.
- Hyysalo, S., Juntunen, J. and Martiskainen, M. (2018). 'Energy Internet forums as acceleration phase transition intermediaries', *Research Policy*, 47: 872-885.
- Janssen, M., Bergek, A. and Wesseling, J. (2022). 'Evaluative systemic innovation and transition programmes. Towards a culture of learning', *PLOS Sustainability and Transformation* 1(3): e0000008.
- Kanda, W., del Río, P., Hjelm, O. and Bienkowska, D. (2019). 'A technological innovation systems approach to analyse the roles of intermediaries in eco-innovation', *Journal of Cleaner Production*, 227: 1136-1148.
- Kanda, W., Huisma, M., Kivimaa, P. and Hjelm, O. (2020). 'Conceptualising the systemic activities of intermediaries in sustainability transitions', *Transitions*, 36: 449-465.
- Kivimaa, P., Kangas, H-L., and Lazarevic, D. (2017). 'Client-oriented evaluation of 'creative destruction' in policy mixes: Finnish policies on building energy efficient transition', *Energy Research & Social Science*, 33: 115-127.
- Kivimaa, P., Boon, W., Hyysalo, S. and Klerkx, L. (2019b). 'Towards a typology of intermediaries in sustainability transitions: a systematic review and a research agenda', *Research Policy*, 48: 1062-1075.
- Loorbach, D. and Rotmans, J. (2010). 'The practice of transition management: Examples and lessons from four distinct cases', *Futures*, 42: 237-246.
- Magro, E. and Wilson, J.R. (2013). 'Complex innovation policy systems: Towards an evaluation mix', *Research Policy*, 42/9: 1647-1656.
- Mickwitz, P., Neij, L., Johansson, M., Benner, M. and Sandin, S. (2021). 'A theory-based approach to evaluations intended to inform transitions toward sustainability', *Evaluation*, 27 (3): 281-306.
- Molas-Gallart, J., Boni, A., Schot, J. and Giachi, S. (2020). *A Formative Approach to the Evaluation of Transformative Innovation Policy*. Transformative Innovation Policy Consortium (TIPC) Research Report, July 2020.

- OECD (2015). *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development*, The Measurement of Scientific, Technological and Innovation Activities, Paris: OECD.
- OECD/Eurostat (2018). *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation, 4th edition*, The Measurement of Scientific, Technological and Innovation Activities, Paris: OECD.
- Ofir, Z. and Rugg, D. (2021). 'Section on International Developments in Evaluation: Transforming Evaluation for Times of Global Transformation', *American Journal of Evaluation*, 42 (1): 47-52.
- Patton, M.Q. (2010). *Developmental evaluation. Applying complexity concepts to enhance innovation and use*. Guilford Press: New York.
- Patton, M.Q. (2016). 'The developmental evaluation mindset: Eight guiding principles'. Patton, M.W., McKegg, K. and Wehipeihana, N. (eds.) *Developmental Evaluation Exemplars: Principles in practice*, pp. 289-312. New York, NY: Guilford.
- Pawson, R. and Tilley, N. (1997). *Realist Evaluation*. SAGE Publications: London.
- Robinson, D., Schoen, A., Larédo, P., Molas Gallart, J., Warnke, P., Kuhlmann, S., Ordóñez-Matamoros, G. (2021). 'Policy lensing of future-oriented strategic intelligence: An experiment connecting foresight with decision making contexts', *Technological Forecasting & Social Change*, 169 (2021) 120803.
- Schot, J. and Geels, F.W. (2008). 'Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy', *Technology Analysis & Strategic Management*, 20: 537-554.
- Schot, J. and Steinmueller, W.E. (2018). 'Three Frames for Innovation Policy: R&D, Systems of Innovation and Transformative Change', *Research Policy*, 47/9: 1554-1567.
- Van Lente, H., Hekkert, M., Smits, R. and van Waveren, B. (2003). 'Roles of Systemic Intermediaries in Transition Processes', *International Journal of Innovation Management*, 7 (3): 1-33.
- Van Mierlo, B., Regeer, B., van Amstel, M., Arkesteijn, M., Beekman, V., Bunders, J., de Cock Buning, T., Elzen, B., Hoes, A-C. and Leeuwis, C. (2010). *Reflexive Monitoring in Action – a guide for monitoring system innovation projects*. Wageningen/Amsterdam: Communication and Innovation Studies, WUR.
- 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*, 41/6: 1037-1047.

Appendix II: Overview of principles and approaches for evaluation of transformative innovation policy or sustainability transitions – excerpts from articles

	Molas-Gallart et al. 2021	Patton 2021	Turnheim et al. 2015
Principles for evaluating transformation	<p>From work of Transformative Innovation Policy Consortia (TIPC), six principles:</p> <ul style="list-style-type: none"> - Adopt a formative approach to evaluation - Integrate evaluation with policy design and implementation - The evaluation process should be inclusive and participatory - Use of mix of methods and techniques - Use a nested approach to assess multi-level TIPs - Use a flexible theory of change, comprised of five elements: <ul style="list-style-type: none"> - Context (background socio-technical landscape influencing regime change but not directly addressed by intervention) - Inputs (resources available to actors to enact change, including inputs from policy intervention) - Activities (interventions which together constitute the experiment) - Transformative Outcomes (in niche building, niche expansion, regime opening – see Table 1) - Impact (emergence of new, sustainable socio-technical systems that deliver on ultimate policy goals – e.g. reduction of inequality, CO2 emissions, air pollution, etc.) 	<p>Adapted from Development Assistance Committee (DAC) Network on development evaluation (OECD/DAC 2019), evaluation criteria for evaluating transformation:</p> <ul style="list-style-type: none"> - Transformation Fidelity Criterion (the extent to which the realities of transformational change initiatives match transformational aspirations and rhetoric) - Complex Systems Framing Criterion (Assess systems transformation using systems thinking principles and complexity concepts) - Eco-Efficient Full-Cost Accounting Criterion (Document and assess the full costs and benefits of systems transformations, including economic, social, and environmental dimensions) - Adaptive Sustainability Criterion (Evaluate transformational sustainability as manifesting ecosystem resilience and adaptability at the nexus between humans and the environment) - Diversity/Equity/Inclusion (DEI) Criterion (Evaluate how transformational engagement manifests the values of diversity, equity, and inclusion together) - Interconnectedness Momentum Criterion (Identify, understand, and evaluate the interconnections that are essential and integral to transformation) 	<p>An integrated evaluation of sustainability transitions should be comprised of the respective strengths of quantitative systems modelling, socio-technical analysis and initiative-based learning:</p> <ul style="list-style-type: none"> - An ability for developing future projections/scenarios: explicit goals based on policy intentions and targets, and an assessment of how we can get from the present to these objectives, informed by: - (focussed) in-depth analysis: an understanding of the recent past and present (the degree of inertia of regime trends, possible alternatives), an understanding of where are we currently heading (niche momentum, regime transformation, etc.), including: - (generalizable) lessons about the scaling of experimentation: an understanding of what is happening on the ground, emerging trends-in-the-making, the determinants of successful implementation and scaling up, etc.
Approaches for evaluating transformation	<p>Instead of providing a tool for comparison across experiments (Luederitz et al. 2016), aim to develop an approach within an experiment or policy...a policy evaluation tool applicable to a specific intervention and providing input into the other policy cycle tasks (policy definition and implementation).</p> <p>Reflexive monitoring and evaluation approach (compared to more common 'result-oriented' evaluations) that considers learning how to contribute to system innovation the central goal of evaluation... a formative evaluation approach conducted with the participation of stakeholders with the main purpose of improving the definition and implementation of the interventions being evaluated</p>		

	<p>Design, adaptation and regular iterations of a flexible theory of change using participatory techniques (to foster learning and reflexivity among participants and to help assess if the policy is contributing to move towards its objectives)</p> <p>Flexible ToC can then be used for developing indicators for transformative outcomes (TOs). Indicators are the result of a reflexive process involving a wide range of project participants, who then use them to inform assessments of the degree to which they are making progress into the desired trajectory of change and to discuss and guide choices on future actions.</p> <p>Indicators cannot be used as an objective measure to compare between EPEs or on which to base funding decisions.</p> <p>Use of 12 TOs to assess in real-time (a form of constructive monitoring) the degree to which the interventions are progressing towards the achievement of long-term systemic goals.</p> <p>Real-time monitoring embedded in the policy process...The reflexive process provided by formative evaluation and the focus on TOs can drive policies towards achieving their transformative goals.</p>		
--	--	--	--