# A FORMATIVE APPROACH TO THE EVALUATION OF TRANSFORMATIVE INNOVATION POLICY

Jordi Molas-Gallart, Alejandra Boni, Johan Schot, Sandro Giachi

# A formative approach to the evaluation of Transformative Innovation Policy

Jordi Molas-Gallart[1], Alejandra Boni[1], Johan Schot[2], Sandro Giachi

([1]INGENIO (CSIC-UPV); [2]Utrecht Centre for Global Challenges, Utrecht University)

## Abstract

Transformative Innovation Policies (TIPs) propose that addressing the key challenges currently facing our societies requires profound changes in current socio-technical systems. To leverage such 'socio-technical transitions' calls for a different, broad mix of research and innovation policies, with particular attention being paid to policy experiments. As TIPs diffuse and gain legitimacy they pose a substantial evaluation challenge: how can we evaluate these policy experiments with a narrow geographical and temporal scope, when the final objective is ambitiously systemic? How can we know whether a specific set of policy experiments is contributing to systemic transformation? Drawing on TIPs principles as developed by and applied in the activities of the Transformative Innovation Policy Consortium (TIPC) and on the concept of transformative outcomes, this article develops an approach to the evaluation of TIPs that is operational and adaptable to different contexts.

Keywords: formative evaluation, sustainability transitions, transformative outcomes, flexible Theories of Change

## 1. Introduction

Transformative Innovation Policies (TIPs) propose that addressing the key challenges currently facing our societies requires profound changes in current socio-technical systems (Weber and Rohracher, 2012, Schot and Steinmueller, 2018). To leverage such 'socio-technical transitions' calls for a different, broad mix of research and innovation policies, with particular attention being paid to policy experiments. Following Schot et al. (2019) we will call these Experimental Policy Engagements (EPEs). This notion is introduced to signal that transitions are complex and long term processes that can be modulated through TIPs but not controlled. Such EPEs aim at establishing niche innovations, and then accelerating and embedding them, through processes of upscaling, replication, circulation and institutionalisation leading to the opening up and unlocking of the current socio-technical regime and its substitution for a new one.

As TIPs diffuse and gain legitimacy they pose a substantial evaluation challenge: how can we evaluate EPEs or sets of EPEs with a narrow geographical and temporal scope, when the final objective is ambitiously systemic? How can we know whether a specific set of EPEs has set ourselves up on the way to systemic transformation? The problem we face can be seen as a specific instance of the common "attribution" challenge posed

by the impact assessment of policies that occur a long way upstream from their intended final objectives, as for instance societal-challenge driven research policies, or local interventions aiming at socio-economic development (Smutylo, 2001): the results of an intervention can be no more than a contribution to (and not a determining cause of) the systemic changes being pursued.

Several evaluation frameworks and approaches have already been developed for sustainable innovation policies (Popper et al., 2017), sustainable transition policies (Taanman, 2014, Luederitz et al., 2017), and responsible research to leverage sustainable transformations (Daedlow et al., 2016), among others. However, TIPs have specific characteristics that influence the design of evaluation methods. TIPs proponents suggest a set of policy features that form part of the specific transformation logic they seek. For instance, policy interventions have to be participatory, aim for second order learning, address societal and environmental challenges, induce system change, and accommodate controversy (Daniels et al., 2020). We argue that these characteristics have to be extended to policy evaluation methods and practices, and have a bearing on the roles of the actors involved in evaluation activities, as well as changes in the organizational routines within which these evaluation practices are inserted.

This article addresses the following questions: how to evaluate transformative innovation policy activities (projects, programs and policies)? Which evaluation methodology and practice would be most suitable and how to implement it? The article draws on the principles of TIPs as developed by Schot and Steinmueller (2018) and applied in the activities of the Transformative Innovation Policy Consortium (TIPC) (http://www.tipconsortium.net/)[1] as well as on a complementary articles that introduce the notion of transformative outcomes (Ghosh et al., Forthcoming). In Section 2, we discuss the demand side for any evaluation of TIPs and the principles that need to inform this process. In section 3, we discuss various approaches available in the literature closely related to the evaluation of TIPs. In section 4 we present our formative evaluation approach using transformative outcomes (presented in Section 4). This approach aims to be operational and adaptable to the different contexts. In Section 5, we illustrate the implications and potential of applying the proposed evaluation strategy with reference to a case of a TIP that, despite its transformative goals, was assessed using existing evaluation approaches. This case is used because it was part of the research process leading to the formulation of our proposed formative evaluation approach with transformative outcomes.

This research process was part of the co-creation journey between a number of Science, Technology and Innovation (STI) agencies and researchers working together in TIPC. Initial versions of our approach were presented and discussed in four TIPC workshops and training activities in 2018 and 2019 and other bilateral interactions involving officers from science and innovation agencies in six different countries (South Africa,

---

[1] The Transformative Innovation Policy Consortium (TIPC) consists of a set of research and policy partners, including innovation and research agencies from Finland, Sweden, Norway, South Africa, and Colombia. The Consortium's key objective is to examine and expand on current policy approaches to assist in solving urgent social, environmental and economic issues. TIPC aims to shape and deliver a new transformative innovation policy framework. During the process, all participants are positioned as active co-researchers and policy co-designers.

Colombia, Mexico, Norway, Finland and Sweden). The case-study was conducted with the Swedish Innovation agency Vinnova (end of 2018-first months of 2019) in order to assess and learn about the added value of our evaluation proposal for their practice. This article contains a tentative proposal based on this journey and a literature review conducted specifically for this paper. This review is limited to publications that aim at evaluating for transformative change, identified through a snow-balling technique. In addition, we bridge these papers with a specific formative evaluation tradition as will be explained in section 3.

## 2. TIPs principles

### 2.1.  TIP as a third frame for research and innovation policy

TIPs starts from a critical consideration of the shortcomings of current approaches to research and innovation policies. The model proposed by Schot and Steinmueller (2018) establishes the existence of three innovation policy frames. In the "first frame", policy is based on a linear understanding of innovation; innovation emerges from a process that starts with the generation of new knowledge through basic and applied research, the further development of such knowledge into new technologies, which when applied generate welfare and growth. Within this frame, policy (and evaluation) objectives can be defined and operationalized by focusing on the quality, nature and mix of R&D inputs and how they shape the excellence, innovativeness and viability of the knowledge system. Incentives for providing inputs such as intellectual property rights can also be seen as an element of the frame.

The innovation systems literature has provided a 'second frame' for innovation policy, stressing that the progression from new knowledge, to new technologies, innovation and growth is far from automatic and does not necessarily move in a single direction. New technological development can spur, for instance, basic and fundamental research and the extent to which new knowledge and technologies will lead to innovation and growth is contingent upon a variety of institutional factors and the linkages among different participants in an innovation system. The focus on systemic failures has provided a different rationale for innovation policies, moving beyond R&D investment levels to the institutional conditions and inter-organizational links that can promote innovation. Yet the type of innovation that is thus promoted, its direction and characteristics are of lesser concern. Within this frame, policy (and evaluation) objectives can be defined and operationalised by focusing on the scope, scale and quality of interactions among various actors in the innovation system, and the skills they need to participate in the interactions. Another policy objective can be to encourage actors to become more entrepreneurial, including the promotion of commercialisation activities among knowledge producers.

The third innovation policy frame focuses attention on addressing societal and environmental challenges through socio-technical system change (which is different from knowledge production and product and process innovation). From this perspective, the direction of innovation, and the connection between the ecological, social and the technological arenas become key concerns. Drawing on the STS perspective, innovation is not seen as just a technological process, but one which has social and ecological features build into the design of each technological trajectory. In other words, technology is not neutral but has specific societal and even moral choices and preferences embedded in its design. This frame calls for working with a broader and

different mix of research and innovation policies that aim for socio-technical transitions, or transformations.

Building on the sustainability transitions literature it is assumed that such socio-technical transitions can be leveraged through EPEs supporting three core transition processes: building and nurturing of niches (or alternative practices), expanding and mainstreaming niches into the wider world (or system diffusion) and the opening up and unlocking of regimes (Grin et al., 2010, Markard et al., 2012, Kivimaa and Kern, 2016, Schot et al., 2019). These EPEs have explicit directionality; that is, they are expected to stimulate socio-technical regime change in more sustainable directions. To achieve their transformational goals, policy interventions have to include the development of transformational policy evaluation methods and practices and imply crucial modifications in the roles of the actors involved in evaluation activities as well as changes in the organizational routines within which evaluation practices are inserted.

## 2.2. Principles of TIP and implications for evaluation practice

TIP explicitly seeks to address social and environmental challenges at the onset, rather than assuming or hoping that economic growth will eventually result in the resolution of relevant social and environmental problems. During the TIPC pilot phase (Sep 2016 - Feb 2017), TIPC studied transformative innovation policies in its member countries through a set of pilot studies (Ghosh et al., Forthcoming). To select the case studies the Consortium developed a set of six criteria identifying the transformative characteristics of an innovation policy: 1) directionality, 2) inclusiveness, 3) societal goal, 4) systemic impact, 5) deep learning and reflexivity, and 6) conflict vs consensus (Schot et al., 2017, Daniels et al., 2020).

Directionality means that innovation policy will not just stimulate specific technological options, but instead look into the social and environmental drivers and consequences of each option and then aim for a deliberation on desirable directions and eventually deliberately and explicitly foster some desired directions for innovation, while blocking undesirable ones. Of course this is an iterative process, and not all consequences and directions can be known upfront, so a flexible and modular approach is required. Inclusiveness refers to the meaningful involvement of a diversity of actors, from civil society to users and local communities, both niche and regime actors as well as groups that are seldom involved in STI policy. Social participation is tightly linked with directionality, since it will shape the debate about directionality.

Connected with these two principles, "Societal goal" states that innovation is not a goal in itself but rather is a tool to address the challenges that our societies face (like, for instance, those encapsulated in the Sustainable Development Goals – SDGs). To achieve these goals, we need profound changes in the socio-technical system, and therefore innovation policies should aim at "systemic impact". To achieve these systemic impacts, TIPs need to incorporate deep learning and reflexivity, which in this context we take it to imply the questioning and reframing of the underlying rules of the dominant socio-technical systems. These rules pertain to all dimensions of the system from science and technology, to industry strategy, to markets and users, government regulations and cultural beliefs. Second-order learning typically emerges if the diversity of opinions and beliefs among stakeholders are acknowledged and embraced. Because focussing on disruptive change can result in disagreements among the stakeholders, TIPs require broad consultation processes to discuss different rationales and perspectives in order to broaden the scope of inputs into policy definition, uncover innovative ideas and minimize legitimation problems later on.

These criteria have important implications for evaluation. Directionality recognizes that TIP will open up a debate about choosing socio-technical pathways, and therefore foster some directions for innovation. Consequently, traditional indicators (employment, new markets and market shares, patents, publications etc…) may not necessarily provide indicators of success. 'Inclusiveness' requires participation and meaningful debate about policy goals, values, and ultimately about the criteria by which such policies should be evaluated. Second order learning implies we need to be open to the reconsideration of basic assumptions, and therefore evaluation criteria, throughout the processes of TIP implementation.

The six criteria are rooted in a specific understanding of transformation. They rest on the Multi-Level Perspective (MLP) on socio-technical transitions as defined by Geels (Geels, 2002, Geels, 2010) and Geels and Schot (2007). This theoretical framework sets the basis for the logic of TIPs, and provides the foundation on which we will build the TIPs "generic Theory of Change".

Some elements of this theoretical body are of particular relevance here. First, transformation requires a change (transition) in socio-technical systems. These systems are stable and dominant configurations of practices, relations, discourses, culture, legislation etc. providing ways of realising a particular societal function (Smith et al., 2010). Second, these system elements are put in place, maintained and destroyed by a wide range of actors whose behaviour is configured by rules while they also construct in concrete actions. Here we draw on the sociological duality of structure principle introduced by (Giddens, 1984). Actors are not passive rule-followers but knowledgeable agents who actively use rules to interpret the world, make decisions and act. These rules contain behavioural instructions, beliefs and values concerning all system dimensions. Together they form a socio-technical regime. In the end a system transformation is not only about changing the system, but also about constructing a new regime (rule-set) using the innovative capacities of all relevant actors. It is for this reason that learning, and reflexivity has such an important role to play. Actors need to use their agency, question the rules they use in their daily practices, unmake them, and become active rule-makers.

Third, a system transition can be understood as the interaction across three levels: landscape, regime and niche. Niches are protective spaces where different ideas, models, configurations and ways of doing try to survive and develop. Niches present configurations whose characteristics are different from those of the regime: they may work with different principles; may use different technologies; present different relations between stakeholders, channels and user practices; or may privilege different sources of knowledge and alternative cultures. Systems and regimes are usually stable, but are permanently exposed to pressures derived from external, powerful and long-term economic, social, cultural or environmental trends (Rotmans et al., 2001), which constitute the "landscape". Instead, niches usually evolve quickly as they are spaces of experimentation and change (Geels, 2002). Third, niches are the place of transformative ideas and practices, but their potential is constrained or enabled through the more powerful structures of the regime (Bos and Grin, 2008). Systems transitions may take place when the regime is destabilised because of the heavy pressure of the landscape, so windows of opportunity may be open for niches—if they are mature enough—to influence or even completely replace the regime (Geels, 2002).

TIPs are based on this logic and will therefore typically combine EPEs aiming at building up and nurturing, expanding and mainstreaming of niches with those aiming at the destabilisation of the dominant unsustainable socio-technical regime, and ultimately

a movement towards a new more sustainable socio-technical regime. They will need to address the different interconnected system elements, and the underlying set of rules sustaining these elements; therefore, not only will they attempt to introduce a specific set of technologies, but they will also deal with the changes needed in rules and regulations, markets, institutions, practices and culture that are required for technological change to be able to be implemented in such a way as to enact regime change, and address current societal challenges.

TIP evaluation needs will refer to this theoretical framework as the main reference to guide the definition of the logic underlining each policy intervention. We will propose and approach that is consistent with TIP principles and informed by transitions theory in section 4. Before, we will discuss existing approaches to the evaluation of policy approaches that share with TIP an ambition to generate system change in order to tackle societal challenges, particularly in the field of environmental sustainability.

## 3. Sustainability and transformation: evaluation challenges and approaches

TIP evaluation involves assessing the changes associated with or leading to socio-technical transitions. For a variety of reasons this is a very challenging task (Kivimaa et al., 2017, p. 118). First, as argued above, we encounter the problem of relating ambitious medium and long term systemic goals with locally bounded EPEs. Yet, TIPs are not alone in facing this problem. There is a set of related policies aiming to enact systemic change to address environmental problems that have already been exposed to evaluation attempts. For instance, several authors have addressed the evaluation of policy mixes directed to system transformation (Kivimaa et al., 2017, Magro and Wilson, 2013, Janssen, 2019), and the evaluation of transitions programmes and experiments (Taanman, 2014, Heiskanen and Matschoss, 2018, Luederitz et al., 2017). There is also relevant work evaluating specific initiatives in support of climate objectives (Hildén et al., 2014), and a broad field of transdisciplinary research on socio-ecological systems that has also generated evaluation efforts (Holzer et al., 2018). There is also significant work focussing on the use of a reflexive approach to evaluation to encourage learning within a group of diverse actors who seek to contribute to system change (Arkesteijn et al., 2015, van Mierlo et al., 2010a, van Mierlo et al., 2010b).

These approaches exemplify the small, distinct and thinly connected evaluation communities that have evolved around policy evaluation in these areas and reveal the use of a considerable variety of methods and techniques, some of which will be relevant for our effort. Thus for instance, Taanman (2014) and Heiskanen and Matschoss (2018) use "Theories of Change", while Magro and Wilson (2013), Luederitz et al. (2017) and Janssen (2019) use formal protocols applied in sequential stages without linking them to an explicit programme or change theory. van Mierlo et al. (2010b) and Arkesteijn et al. (2015) propose a "reflexive monitoring in action" methodology with the scope of fostering recurrent reflection on 1) the relationships between project activities and results and its institutional settings and 2) the ambition to change in both short term-actions and long-term goals and future perspectives.

Significantly, the studies also differ on the function they attribute to evaluation. Taanman (2014) and Heiskanen and Matschoss (2018) explicitly advocate formative evaluations; that is, evaluations conducted in close collaboration with stakeholders with a view to improve policy design and implementation. Magro and Wilson (2013), Luederitz et al. (2017), Janssen (2019) and Holzer et al. (2018) add accountability to the

learning and improvement goals that characterise formative evaluation, placing varying importance on the accountability function. van Mierlo et al. (2010a) and Arkesteijn et al. (2015) stress how evaluation can stimulate social learning within institutions in order to develop new practices to enact systemic change.

Whatever the differences, all approaches are confronted with the systemic character of the policy aims, and the multi-level nature of the potential policy interventions. 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, eventually, the systemic impacts of policy mixes (Arnold, 2004, Edler et al., 2008). Typically, evaluation efforts in the sustainability transitions area aim at developing frameworks to address the macro-, or meso-level of socio-technical change. These contributions, however, offer little help to evaluate the contribution of local interventions to systemic transformation.

Turnheim et al. (2015) propose to tackle this gap between a specific action and systematic transformation through 'an integration strategy based on alignment, bridging and iteration' of learning-based evaluation of local initiatives with socio-technical analysis at 'regime' level, and quantitative system modelling at the 'landscape' levels. Yet, their proposal is very complex and articulated around the idea of aligning interventions that are inherently different as they operate at different policy levels.

The assessment of multi-level policy action has also been addressed by Taanman (2014) who identified three levels for sustainability transitions: projects, programs and transition field.[2] To assess the way in which "transition programmes" contributed to their ultimate policy goals, Taanman uses concepts from the sustainability transitions literature: transition scenarios, changes in organization, practices and culture, and sustainability criteria. These are important building blocks in the logic of transitions policies and are also stressed by other authors in their evaluation frameworks: Kivimaa et al. (2017) focus on how specific actions lead to changes in institutions and organizations, while Hildén et al. (2014), Luederitz et al. (2017), Heiskanen and Matschoss (2018), and Janssen (2019) define specific dimensions related to the transformation of socio-technical systems that should be shaped by specific actions.

The complexity of the systemic change that the policies are attempting to enact has often drawn evaluators towards a focus on policy processes as a way to disentangle complexity and draw lessons from the evaluation effort. For instance, Taanman (2014) focuses on following up the whole policy process from the very beginning employing a monitoring approach open to experimentation with different evaluation strategies and methods. Similarly, van Mierlo et al. (2010b) and Arkesteijn et al. (2015) suggest that a reflexive evaluation approach should be understood as a "reflexive monitoring in action" considering monitoring as an integral part of the change initiative since its start.

---

[2] Heiskanen, E. and Matschoss, K. (2018) 'Evaluating Climate Governance Experiments', in Turnheim, B., Berkhout, F. & Kivimaa, P. (eds.) *Innovating Climate Governance: Moving Beyond Experiments*. Cambridge: Cambridge University Press, pp. 182-200. and Holzer, J. M., Carmon, N. and Orenstein, D. E. (2018) 'A methodology for evaluating transdisciplinary research on coupled socio-ecological systems', *Ecological Indicators,* 85, pp. 808-819. also carry out their evaluation at different levels, but do not explicitly formalise a multi-level approach.

Focusing on processes requires attention to be paid to the diversity of participating actors. Further, it is common for such actors to be considered an active participant in the evaluation rather than just an evaluation subject. Different approaches organise participation in different ways and seeking different goals. Janssen (2019) highlights the opportunity to involve private firms to strengthen existing knowledge networks, while Kivimaa et al. (2017) presents an evaluative research design where stakeholders (in this case, intermediary companies) are used as the focal point for assessing different perspectives on policy implementation and impact. Luederitz et al. (2017) point out that inclusion requires, beyond mere participation, the creation of opportunities and the empowerment of excluded communities, but they do not provide details as to how this can be done and how evaluators can avoid reproducing the existing structure of power and interests. Heiskanen and Matschoss (2018) argue that the performance of transition experiments can be influenced by participants' perspectives and beliefs, and that evaluators should act as facilitators, not only managing conflicts of interest and power, but also different cognitive and cultural perspectives, and foreseeing and addressing emerging changes and trends, as well as new perspectives going beyond the initial boundaries of the experiment. van Mierlo et al. (2010b) and Arkesteijn et al. (2015) propose a "reflexive monitor", either an external person or someone from the project team, who is observing how the ambition of change is articulated, whether learning is taking place and whether collaborative actions are being designed and carried out.

In short, the variety of perspectives described above have focused on process evaluation, have tended to favour formative and inclusive approaches to evaluation, but have encountered difficulties when trying to implement a multi-level approach linking specific actions with systemic transformation. This remains the most difficult challenge facing the evaluation of TIPs: to establish whether and how upstream policies are contributing to systemic changes which can only occur as the result of complex processes combining many other contributory factors. Some of the proposed evaluation schemes in the area of sustainability transitions respond with complex and ambitious approaches attempting to track the whole pathway linking different sets of policy contributions (emanating from interdependent policy mixes) to instances of changes that reflect systemic shifts. Although theoretically well-grounded, the complexity of these approaches makes it difficult to implement.

Yet, evaluation practice in other policy fields has also encountered the problem of interventions occurring a long way upstream from the desired policy results. Evaluators have developed different approaches focusing policy evaluation on the policy outcomes rather than their long-term impacts (Wilson-Grau and Britt, 2013, Earl et al., 2001, Belcher et al., 2020). Outcomes have been defined "as changes in the behaviour, relationships, activities, or actions of the people, groups, and organizations with whom the programme works directly. These outcomes can be logically linked to a program's activities, although they are not necessarily caused by them" (Earl et al., 2001, p. 1).[3] This definition, which we are going to adopt, focuses therefore on observable changes we are calling transformative outcomes that can be directly linked to the activities originating from an intervention, without requiring that we formally demonstrate that

---

[3] Different outcome-focused evaluation approaches define "outcomes" in different ways. For instance, Belcher et al. Belcher, B. M., Davel, R. and Claus, R. (2020) 'A refined method for theory-based evaluation of the societal impacts of research', *MethodsX*, 7, pp. 100788. consider them only in terms of behaviour change, "conceptualized (and emphasized) as changes in knowledge, attitudes, skills, and/or relationships (KASR)".

they have been caused by them. What needs to be shown is a logical and demonstrable connection between these transformative outcomes changes and a set of activities (Ghosh et al., Forthcoming).

Outcomes are distinct from outputs: the direct results of an activity supported by the intervention. Such outputs are directly caused by the intervention, buy they stop short of representing as they still need to be adopted, or used, or further developed through relevant changes in behaviours, activities, relationships or actions. Here we can make a direct connection with the MLP notion of changes in rule-sets. In other words, to ascertain than an activity is moving its beneficiaries in the directions that the policy designers expect we have to see changes in rules used by actors to enact system transformation. For instance, a sustainable innovation programme can lead to the development of a new energy generation technology reflected in the granting of a patent or an article (outputs). Yet, it is not until such innovation leads to changes in behaviours, beliefs and values among the targeted population (niche and regime actors) that we can say we are on track to achieve the final impacts sought (for instance, the establishment of a new, sustainable, socio-technical system). Such systemic changes are likely to take place well after an intervention has concluded, but for evaluation to assist in programme implementation it has to be conducted while this is still ongoing. Therefore, we suggest that the evaluation of TIPs focuses on the analysis of transformative outcomes than can be expected to accrue while the intervention is still ongoing. Focusing on outcomes is common practice when evaluating policies whose ultimate intended effect will take place a long way "downstream". These outcomes may be traced to the immediate results (outputs) generated by the intervention, but more importantly they are contributing to the systemic change that constitutes the ultimate policy goal. Such links constitute a "Theory of Change" understood as an account of what is expected to happen; that is how policy inputs lead to activities that contribute to relevant changes (Transformative Outcomes) which in turn will contribute to systemic change.

For building this Theory of Change we propose to work with existing transitions theory, in particular the multi-level perspective (MLP). Our approach is similar to theory-oriented approaches (Stame, 2004) in that it revolves around a description of an expected process of change that "considers programmes in their context, which includes actors' environments…and public service culture and behaviour". But unlike most theory-oriented approaches in which the evaluator "interprets" the understanding of what may happen offered by actors involved in the intervention (Stame, 2004), we actively use transitions theory to co-produce with the policy actors a theory of change that focuses on transformative changes.

In the following section we offer a multi-level approach to TIP evaluation. In line with TIP principles and existing evaluation practice we adopt a formative approach focusing on the analysis of transformative outcomes with the objective of improving the definition and implementation of TIPs through participatory evaluation techniques.

## 4. A formative evaluation approach for TIP

### 4.1. Six guiding principles for TIP evaluation

The evaluation approach should reflect the principles of the policy being assessed. Based on the six TIP dimensions and the extant evaluation practice in the area of

sustainable transition policies, we propose a set of six principles to guide the evaluation of TIPs.

1) *Adopt a formative approach to evaluation.* By a formative approach we mean a style of evaluation which is conducted with the participation of stakeholders with the main purpose of improving the definition and implementation of the interventions being evaluated. Under this perspective, evaluation should be understood as a reflexive practice aiming at helping policy actors to navigate their TIPs and contributing to their capacities to do so. In such a practice, failure should be seen as a learning opportunity on the context, conditions and activities conducive to transformation processes. In addition, evaluation can help refining transformative innovation theory by providing information about different transformative outcomes triggered by different policies under different landscape and regime conditions.

Our approach to the application of formative evaluation to innovation policies draws on a stream of evaluation work dating back to, at least, the mid-80s and the evaluation of the U.K. Alvey program (1984 -1990). This evaluation of a British programme to support R&D in the information technology sector, developed a real-time evaluation approach and provides an early example of formative STI policy evaluation (Molas-Gallart and Davies, 2006). This evaluation became a referent in the early 90s and argued that that real-time evaluation had several advantages over ex-post approaches, particularly the fact that it provided actionable feedback to those working in an intervention (Hobday, 1988). The use of evaluation approaches that were explicitly characterised as real-time and formative can be traced to another evaluation of a programme in the IT sector (Eschenbach et al., 1995). Formative approaches were soon after developed as part of a new mode of evaluation in which evaluators would get directly involved in learning exercises with all programme stakeholders, playing more the role of facilitators rather than that of external experts, and leading to a more flexible and experimental approach to innovation policy formulation (Kuhlmann, 1999).

2) *Integrate evaluation with policy design and implementation.* Following from our understanding of formative evaluation, we see evaluation as part of the policy process and, therefore, as a task that should share in the overall characteristics we aim this process to have. Specific policies, their implementation and evaluation should be coherent with the stated research and innovation policy objectives (directionality, societal goals and system impact). Evaluation thus becomes a strategic part of the design and implementation process of TIPs.

3) *The evaluation process should be inclusive and participatory.* The inclusivity characterising TIPs should also be present in the evaluation process. Traditional evaluations are often led by external evaluation experts who implement and plan them. In contrast, participants in TIPs should also join in their evaluation, with external evaluation experts mainly acting as facilitators paying, for instance, attention to the power dynamics that may lead to some voices being heard more than others. Therefore, evaluation should facilitate participation and open debate, channelling power conflicts, and differences in interest and perceptions. The groups and communities participating in the evaluation process will be varied and have different access to resources and, even, different interests. Managers and grassroots participants, for instance, may have different perspectives on the definition of the problems to be addressed, and be unequal in terms of the power they hold. An evaluation design should be attentive to such differences.

4) *Use a mix of methods and techniques.* Rather than being driven by formalised standard protocols, evaluation practice needs to be adaptable and flexible, selecting different methods and techniques according to the policy context and its transformative nature. Quantitative techniques can provide synthetic assessments that allow for comparison across different units of assessment, and can provide, under specific conditions, robust assessments of the net impact of an intervention. Yet, the assessment of social and environmental impact is difficult to be achieved with 'standard' indicators. This difficulty is in part attributable to the nature of social values, which are often linked to incommensurable dimensions and perceived differently depending on cultural background and personal preferences. In these situations, qualitative methods can provide a better approximation to impact assessment by providing a fined-grained, contextualised description of transformative outcomes through detailed narratives. Finally, participatory techniques can help increase participation and inclusiveness of the evaluation process.

5) *Use a nested approach to assess multi-level TIPs.* TIPs can operate at different levels. Niche projects are local initiatives attempting to generate or support a specific niche. Programmes may bring together several niche projects and will seek to develop links and relationships among them that will facilitate scaling up. Finally, several programmes can combine with other policy in policy-mixes which aim to realise socio-technical system change. Impact understood as transformation of a socio-technical system cannot accrue from a single niche-level experimental policy. Each small-scale experiment can contribute to socio-technical change, and such contribution can be enhanced by combining them with those of other experiments grouped in policy programmes. Each policy level can be evaluated, but our expectations of what the policy can be achieve will differ according to its level, and yet such expectations should be coherent across levels. The outcomes that are pursued at project level form part and contribute towards outcomes and impacts pursued at programme and policy-mix level. The outcomes achieved at each level achievement are nested within and will contribute towards those of the higher level. This view leads to a nested approach to evaluation that will refer to the generic "Theory of Change" described below.

6) *Use a flexible Theory of Change (ToC).* Many evaluations use "Theories of Change" to structure their work. The Theory of Change approach to policy evaluation has a long history. The term was coined by Carol Weiss who recommended to identify potential causal models of the programmes, and defined ToCs as 'the chain of assumptions explaining how activities lead step by step to the expected outcomes' (Weiss, 1998, p. 2). Building on Weiss, Connell and Kubisch (1998) define a ToC approach as a systematic and cumulative study of the links between activities, outcomes, and contexts of the initiative. We do not understand these links as a simple cause-and effect relationship, since contexts, activities and outcomes are co-evolving. A ToC is typically defined by policy stakeholders and starts by identifying the main changes that an intervention is aiming to achieve. Policy goals are therefore defined as changes to a baseline situation. Next, participants work backwards from such intended changes to identify the processes that will lead to them, and how these processes will be triggered by the intervention. In this way stakeholders, with the help of evaluation experts, produce an expected process linking the activities[4] triggered by an intervention with its

---

[4] Such activities are the direct consequence of the inputs that an intervention brings to bear. ToC practitioners often use the term "structure" to refer to the wider set of resources that can be called upon to achieve project goals Cohen, S. and Franco, R. (1992) *Rationalising social policy: evaluation and*

results. Our ToCs will be *flexible,* implying that they should not be understood as a fixed causal chain between inputs, activities, transformative outcomes and impacts, but rather they can be revisited and redefined as a result of the formative evaluation process. The ToCs will be used to foster learning (first and, specially, second order[5]) and reflexivity among participants and to help asses if the policy is contributing to move towards its objectives. Following ToC conventions, we will distinguish five elements, which we will align with the Multi-Level Perspective in transitions theory.[6]

- Context: the background "socio-technical landscape" influencing socio-technical regime change, but which is not directly addressed by the intervention
- Inputs: the resources available to actors to enact change, including the inputs provided by the policy intervention
- Activities: the interventions which together constitute the experiment. These activities are linked to:
- Transformative outcomes in three areas drawing upon MLP: 1) Building and nurturing of niches; 2) expanding and mainstreaming niches; 3)opening up and unlocking regimes. All these outcomes are identifiable in individuals, groups and organisations involved in the experiment, they relate to the change of rule-sets.
- Impact: the emergence of a new, sustainable socio-technical system(s) that will deliver on the ultimate policy goals in terms of reduction of inequality, $CO_2$ reduction, air pollution etc.

The focus on transformative outcomes is a key element in our method and is linked to our theoretical understanding of TIPs (Schot et al., 2019). We argue that there are three main transformative processes in the transition from a local niche where new sustainable socio-technical environment emerges to the change in socio-technical regime: (1) Building or constructing the niches; (2) accelerating their growth and expansion and embedding them in the regime, and (3) opening up the existing regime, destabilising their practices and unlocking path dependencies. The three groups of transformative outcomes mentioned above mirror these three processes. Schot et al. (2019) have identified and defined in detail 12 different types of transformative outcomes, four in each of these 3 groups. A summary of this typology can be found below in Table 1. When co-constructing the ToCs with experiment participants we will identify how the expected outcomes can be mapped against these 12 types.

---

*viability.* Santiago: CEPAL.. Structural elements include the budget and the human resources available to support a programme, its planning, method and operational principles, and the norms under which the intervention operates.

[5] The possibility of reviewing and changing the ToC is a natural consequence of our emphasis on second-order learning as one of the objectives of the evaluation process. Second order learning implies the revision of original assumptions, values and objectives and therefore a possible redrafting of the original ToC.

[6] The distinction between outputs, outcomes and impacts is very common in evaluation, although the specific definitions can vary somewhat. For instance, Boekholt et al. (2014) define outputs as the direct results generated by the intervention, outcomes as the immediate benefits for the beneficiaries of the intervention, and use the term "impact" to refer to the wider effects on society.

*Table 1 Twelve types of transformative outcomes, adapted from Schot et al. (2019), 2019; Ghosh et al. (Forthcoming)  Ghosh et al, forthcoming*

| Niche building | |
|---|---|
| Shielding | Offering protection for niche experiments and normalizing these protection measures. Protection can be offered through subsidies but also market benefits, such as a VAT exemption, or cultural protection by trying to change the meaning or perceptions of a specific solution through a media campaign. |
| Learning | First order (optimising existing behaviour) and second order  (changes in frames and assumptions) in or across several system dimensions (science, technology,  innovation; markets; culture & symbolic meanings; industrial strategy) |
| Networking | Participation in the niche of a wide range of diverse stakeholders Diversity in terms of niche and regime actors, and in terms of regime dimensions<br><br>Building and strengthening ties among actors in a niche<br><br>Creation of a community of practice ensuring resource mobilisation<br><br>Emergence of intermediaries in facilitating the above |
| Navigating expectations | Creating space for voicing new and alternative expectations and bridging the diversity of expectations building a shared vision. |
| **Niche expansion and embedding** | |
| Upscaling–Increasing user adoption | Spread of the adoption of new practices and rules, bandwagon  effect. |
| Replication | Replication of niche conditions in different contexts<br><br>Adaptation of a niche in a different locality |
| Circulation | Circulation of ideas, people, tacit knowledge, rules across niches and system dimensions<br>Emergence of system intermediaries |
| Institutionalization (formal and informal rules) | Developing standard definitions, narratives, regulations and preferred types of behaviours, beliefs and values<br><br>Establishment of certification schemes, protocols…<br><br>Development of a mature market niche |
| **Opening up and unlocking regimes** | |
| Destabilising and de-aligning regimes | Disrupt policy frameworks and governance arrangements taking advantage of tensions between regime dimensions<br><br>Phasing out of policies and implementation of other policies disrupting the dominant  socio-technical system |
| Unlearning and deep learning of  regime actors | Second order learning among regime actors - change  existing values and beliefs<br><br>Unlearning routines based on existing skills and capabilities<br><br>Emergence of new policy assumptions |
| Empowering niche-regime interactions | Creation of formal and informal linkages between niche and regime actors<br>Emergence of intermediators facilitating such linkages |
| Changing perceptions of landscape pressures | Regime actors develop new interpretations of the nature and consequences  of trends (such as climate change, loss of biodiversity, pollution, rising inequality, digitalization, urbanisation) and shocks |

## 4.2.    Building ToCs for TIPs´

### 4.2.1 A generic ToC

As described in section 3, Transformative Innovation Policies are informed by transitions theory and, in particular its multi-level perspective. Although the specific ToC that will be built for each project, programme or intervention will be designed by project participants to address the specific circumstances in which the policy is operating and its specific objectives, not all participatory policy designs will lead to TIPs. In the approach, we advance that a TIP needs to respond to the characteristics detailed in section 3 and, among them, be in conversation with the main tenets of the Multi-Level Perspective to socio-technical transitions. In other words, MLP provides a generic ToC that acts as a keystone for the initial definition of specific (project or programme level) flexible ToCs. We do expect and welcome that in this process specific choices are made, and participants will blend and overlay their own implicit or explicit ToC with one in which transformative outcomes are given a central role.

Our generic ToC first states that addressing current societal challenges like climate change or the Sustainable Development Goals requires a change of socio-technical system. Processes leading to such changes have three components: niche building, expansion and de-stabilisation of regimes, which all can be developed through EPEs. If successful, the EPEs will turn into drivers of transformation, and altering the socio-technical system into new configurations that respond to the societal challenges being addressed.

### 4.2.2. A three-step sequence to build specific ToCs

The generic ToC provides a frame to understand the location and ultimate logic of specific EPEs. Yet there is a lot of variation in potential EPEs, and they need to defined with their specific objectives and immediate context and challenges in hand. TIP relies on a set of interrelated EPEs operating at different levels: different STI projects often defining local experiments operating in local niches, and strategies to deepen and scale up the niches, by combining related sets of projects grouped in STI programmes, or, more ambitiously, combining different sets of policies, not only STI ones, but also sectoral ones, such as energy, transport, healthcare, and food policies. Such a policy mix would allow to target all system dimensions, including ones related to the market structure, governance, culture and industry structure. The processes, timelines and immediate objectives of each initiative are different. TIP policy requires the definition and implementation of specific policies at each level of intervention and of the interactions among them. STI projects and programmes can be enacted by dedicated ministries and/or government innovation agencies, and research councils, or by sectoral ministries or agencies who have incorporated a STI function, for example medical STI activities are often carried out by dedicated actors. Each specific ToC will define the inputs (resources), activities and outcomes of a specific intervention taking into account its level.

We propose three steps to develop a specific ToC.

1) Identify the level of the TI experiment (EPE) and the main actors involved. We distinguish three main levels:

- Projects defined as experiments to stimulate transformations and address social and ecological challenges. .

- Programmes linking several of such projects to enhance synergies and generate a wider scale & scope for working on transformations.
- Policy mixes constituted bringing together science, technology and innovation (STI) projects and programmes with other sectoral policies targeting different dimensions of specific socio-technical systems.

2) Identify the key evaluative dimensions to be assessed. These will be different depending on the levels above: project, programme or policy-mix.

At the project level, we distinguish:

- The inputs needed to carry out the project (human, financial, material, organizational, etc.).
- The activities the project aims to carry out.
- Outcomes: the main transformative outcomes generated directly by the STI experiment, typically  we expect these to focus more on shielding, learning, networking, and expectation dynamics thus Niche Building Transformative Outcomes, but others may well come into view (see Table 1 above)
- Impacts: projects are the first step in a long chain of events leading to systemic impacts in the form of socio-technical transitions, but, typically, a specific project in itself cannot cause such impact; therefore, project-level evaluations will not normally aim at identifying systemic impacts.

At the programme level, we distinguish the following:

- The inputs needed to carry out the programme (human, financial, material, legal, organizational, etc.) including the inputs invested in all the specific projects conforming the programme and those required for programme level activities (management, coordination, niche protection and empowering, etc.)
- The activities the programme aims to carry out. These include not only a wide range of projects, but also activities at the programme level to enhance synergies.
- Outcomes: at the programme level the transformative outcomes observed will be the sum of those generated by the projects and programme activities. We expect that STI programmes are in a position to enable more circulation, replication, upscaling and institutionalisation, thus niche expansion and embedding" transformative outcomes. (see above Table 1).
- Impacts: akin to the project level, STI programmes on their own will seldom generate changes in socio-technical regimes within the scope and timeframe of a programme.

At the policy-mix level we distinguish:

- All the inputs (resources) invested to achieve a socio-technical transition, including those invested through STI programmes and those invested in accompanying, complementary sectoral policies.
- How the different programmes and policies interact among each other and complement other initiatives oriented to protecting, nurturing and empowering the project experiments (including changes in regulations, institutional adaptations, etc.)
- Outcomes: in addition to the outcomes generated by the policies and programmes integrating the policy mix, we would expect the combination and interaction among transformative programmes to contribute to all outcomes, but specifically also to generate outcomes related to the opening up and unlocking of regimes (disruption of policy framework and governance arrangements in the dominant regime, second-

order learning among their actors, etc.), therefore shaping communities and groups outside those directly involved at the STI project and programme levels

- Impacts: emergence of new, sustainable socio-technical regimes that can be traced back to activities generated by the policy-mix under assessment. The new regimes will be characterised by substantial reductions in the use of fossil fuels, reductions in emissions, new industrial structures and patterns of consumption supported by changed cultural practices and assumptions.

3) Define a specific ToC consistent with the generic ToC, specifying the pathways that are expected to link the different dimensions defined above; i.e., how the inputs and activities are expected to bring about transformative outcomes. This is a creative process in which context specific TOC aspects are reinterpreted and reframed through drawing on the generic ToC. The main aim of this exercise is to build a ToC aiming at transformation, and by doing so overcoming the transformational failure of many STI policies. Ultimately, we may also want to link outcomes to impacts, e.g. how all outcomes will be combined to help generate a change in various socio-technical systems. For this step we may need new theory and method development, for example using socio-technical scenarios and Deep Transition thinking (Schot and Kanger, 2018). Each participant may implicitly or explicitly hold a different ToC in mind. From a formative evaluation perspective, one of the roles of the evaluators is to enable and facilitate a discussion on the different interpretations, and reach a broad-based agreement on an initial ToC than can be modified during the evaluation process. The definition of ToCs and their use in evaluation should involve a broad range of stakeholders, and their selection is an important part of the evaluation process: selected stakeholders should have legitimacy to participate in the evaluation process and the evaluation process, in order to be useful, should be sensitive to the different needs and values of the participants.[7]

## 5. The implications of our approach

In this section we provide a brief illustration of the differential implications of applying our approach to a transformative innovation programme that had already been implemented and assessed. The case is a long-term project on circular economy in the emerging Forest Chemistry sector, funded by the Swedish Innovation Agency (Vinnova) Challenge-Driven Innovation (CDI) programme. The case was selected together with the partner, with the objective of illustrating the value-added of our approach, particularly by comparing it to the evaluations of the programme that had been undertaken and considering the results achieved by the programme under transformative lenses.

We first present the characteristics of the CDI programme in the context of Sweden's needs for transforming the industry and the business sector in a sustainable way and present the specific project with its transformative goals. We then compare our approach to the archetypical evaluation approach that was used for the project and

---

[7] 'Responsiveness' Abma, T. A. (2005) 'Responsive evaluation: Its meaning and special contribution to health promotion', *Evaluation and Program Planning,* 28, pp. 279-289. is a term that has been used to refer to evaluations that are sensible to the needs of the different actors involved into the policy intervention.

highlight how our formative evaluation framework could enhance socio-technical transformation.

## 5.1.    The Swedish Challenge-Driven Innovation programme

The origin of the CDI programme lies in the 2009 Lund Declaration and the need to move the Swedish STI system towards flexible approaches able to tackle current societal challenges. Vinnova launched the programme in early 2011 with the main purpose of converting societal and environmental challenges into opportunities for economic growth. CDI focused on four related areas: Information Society 3.0, Sustainable Attractive Cities, Future Healthcare, and Competitive Production. "These were all areas in which Sweden had both a strategic interest and a good innovation track record" and where "advances in the development of adequate solutions to many societal problems will need to be made" (OECD, 2016).

The impact goals of the CDI programme aimed at generating solutions for both sustainable growth and internationalisation of Swedish technology. In operational terms, the programme sought to improve coordination and mobilisation amongst business actors, and promote cross-sector collaboration and user- and demand-driven innovation initiatives. Each project had to illustrate its impact logic according addressing specific challenges in line with the programme's objectives. The CDI programme funded projects following a stage-gate process based on three stages: initiation, collaboration and follow-up. The projects were thus assessed before being allowed funding for scaling up in the subsequent stage.

The CDI programme displayed several of the characteristics of a TIP. For instance, Smallman argued that it had "[…] generated a shift in the funding organisation, away from an unspoken focus on technical innovations to a much broader concept of innovation" and that "[…] working practices at Vinnova have also changed as a result of the programme. The range of stakeholders who receive funding from Vinnova has widened, dialogue and collaboration between officers in various departments has increased and the organisation has taken up the important focus on societal challenges" (Smallman, 2018, p. 250). Fuenfschilling et al. (2017) performed a review of the CDI programme under the lens of transformative change and highlighted that it showed several transformational features such as the focus on innovation beyond technology (e.g. social and organisational innovation), on solving social and environmental problems, on the inclusion of new types of actors, and on learning. Despite the experimental nature of the programme, these scholars also highlighted that the current traditional evaluation schemes available are suited neither to support constant reflection and deep learning from the activities nor to track progresses towards socio-technical transformation. The overall impression is that the projects addressed niche-building processes more than accelerating and embedding niche innovations or destabilising existing socio-technical regimes. The involvement of different types of partners in the projects is likely to be recognised as a success of the CDI programme, especially for companies and research institutes, but the inclusion of other actors (universities, users) could have been more intense, especially in terms of systemic integration (OECD, 2016).

We can hypothesize that the failure of the programme to progress towards the embedding and acceleration of transformative innovations could have been, at least, partially addressed had the evaluation processes adopted the formative approach suggested here, instead of the traditional summative methods required to implement a stage-gate approach. To illustrate how this could have been done, we use the case of the

Forest Chemistry project funded by the CDI programme, already reported by Fuenfschilling et al. (2017).

## 5.2. The Forest Chemistry project: an innovation journey

'Forest Chemistry' is a project funded by Vinnova's CDI Programme within the area 'Competitive Production'. It took place between 2011 to 2017, and developed through the three-stages that the CDI programme envisaged: the Forest Chemistry (*Skogskemi*) projects I and II, and the Forest Methanol (*Skogmetanol*) project. The aim of the Forest Chemistry projects was developing new technologies for producing 'green chemicals' using the residues of the purification of raw materials by the forestry industry firms. The main technological objective was the development of a system by which the methanol generated by sulphate pulp mills dedicated to paper production is used for purifying NOx emissions in the local chemical industry. Underpinning this circular economy model there is a vision of socio-technical transformation: there was a need to establish new links between paper pulp mills and the chemical industry and such links require interventions at different levels. For instance, government subsidies would be required for the price of the methanol supplied to the chemical industry to be attractive.

The systemic nature of the endeavour is exemplified by the variety of participants in the project and the need to establish new links among them. Participants included sulphate pulp mills, chemical companies and a 'Support Platform' formed by Vinnova policymakers, researchers, local associations of green activists, workers from the companies, and representatives of the forest and chemical industries.

The participants evolved through different project stages:

- In Stage I, the project included only the leader organisation, Processum – Biorefinery development, a subsidiary organisation of the RISE Research Institute of Sweden group.
- In Stage II, it included a wide network of organisations; mainly, publicly owned companies from the chemistry cluster of the Stenungsunds region (western Sweden) and the forest industry from north eastern Sweden, and also public organisations including universities.
- In Stage III, the network was reduced to a small group of public companies and institutes who decided to continue with the experiment.

The project started with a preliminary study on the components present in forest raw material and its potential use in the chemical industry. Based on its results, RISE Processum pursued increased co-operation between forest and chemical industries to identify chemicals, processes and value chains with large potential. Therefore, in the first stage (November 2011 – Mar 2012) the project focused on knowledge generation, identifying three value chains with the greatest technological and market potential for Sweden: methanol, butanol and olefins. In Stage 2 (Aug 2012 – Nov 2014), pre-feed, system analyses and technical evaluations were performed leading to a focus on the development of technology for cleaning the methanol from stripper gases. A new company joined the consortium in this phase: a sulphate pulp mill that had a high emission of NOx gases, and therefore a need for an improved cleaning process. In the third stage (May 2015 – Jun 2017) the project reached maturation. In 2015, the consortium included five actors operating in different stages of the value chain for cleaning stripper gases in methanol: research organisations, equipment suppliers,

sulphate pulp mill and an organisation representing end customers. In 2016, the consortium achieved the demonstration of a flexible pilot equipment designed by the equipment supplier and installed on the sulphate mill. A 1,000 hours of continuous operation test showed that the pilot equipment obtained results in line with the lab environment and that the purification process worked (NOx emissions were significantly reduced, almost enough to use the methanol as a green input chemical). However, in 2017 the utility equipment manufacturer decided to sell the technology.

According to Vinnova representatives, the project suffered from a series of obstacles and bottlenecks that are common in this type of projects and partnerships. There were coordination difficulties and lack of communication among subprojects and partners. Partners from different sectors did not reach a common understanding on the long-term tasks and goals of the project. Therefore, although short-term commitment to the project was successful, the project could not ensure the long-term engagement of participants.

## 5.3. The evaluation approach

Despite its experimental nature, the Forest Chemistry project followed a linear approach from design to implementation, with limited opportunities to diversify project options and redefine arrangements. Broad and deep expectations played a key role in the first stages of the projects (Allmér, 2017), but stage II was critical for project development as participants closed down alternative options for technology development. In so doing, techno-economic viability was the key evaluative criterion instead of other criteria linked to the social and environmental challenges faced. Broadening and deepening system networks was a concern throughout the project, as well as some form of learning, although second-order learning and reflexivity were not explicitly addressed. In short, although the Forest Chemistry project aimed at transformative change and displayed some of the TIPC features, there were key elements of its implementation that separated from the TIPC model. Paramount among them was the way in which evaluation was implemented.

The stage-gate mechanism led to a summative approach accompanies by tight timeframes, which many participants criticised: it encroached with the long processes of dealing with the regulatory and technical aspects of developing applications and hindered the broader involvement of civil society actors (Fuenfschilling et al., 2017). As a result, the easiest and more straightforward options tended to be chosen instead of pursuing riskier and potentially transformative alternatives.

Another problem was the existence of different and, sometimes, competing expectations and visions among partners. Allmér (2017) argues that the support the project gave for options with the most credible expectations of financial benefit and scaling up potential, was opposed by the chemical companies. To an extent, this occurs because of defective feedback mechanism during project implementation leading to choice that were not consistent with initial expectations and visions, and left no scope for questioning, in a participatory and reflexive way, the assumptions behind the choices made.

A formative evaluation based on the participatory design and adaptation of a ToC consistent with TIPC principles would have helped address these gaps. A formative approach permits to manage expectations and visions of the emerging network since the very beginning and to express them through a flexible ToC, designed and adapted by means of participatory techniques. This ToC can then be used for developing indicators; but it must be noted that these indicators are the result of a reflexive process in which the goals and the process leading to them are defined first, and from them the

participants define the properties that the project aims to change. Observable indicators that adequately reflect this processes need then to be identified. Note that this process is very different from the requirement to find easily quantifiable and difficult to "game" indicators, which can also allow a comparative measure (usually against a benchmark) of project achievements. These latter indicators are needed to make "stop-go" decision at the stage gate, and can be typically found in measures of technical or financial performance. Yet, transformative policies need to be guided by more complex socio-technical achievements which can seldom be rendered by easily measurable and comparable indicators. In a formative approach the indicators, linked to the ToC, will be use to inform assessment by the project participants of the degree to which they are making progress into the desired trajectory of change (instead of being used by external actors as an objective measure on which to base funding decisions).

If we look into the Forest Chemistry project goals and context, we can draw a simple ToC that could have been used as the foundation for the development of the specific ToC by project participants. This transformative ToC would have integrated the participants' narrative for change within the broader ToC of the CDI programme and its vision for socio-technical transformation. It would have established expected interconnections among a set of elements like the following.

- Context:
    - Social and political pressure on use of fossil fuels given climate change, pollution and sustainability concerns
    - SDGs (e.g. 9 Industry and infrastructure; 12 Responsible production and consumption, 13 Climate action) and Swedish strategy on SDGs
    - Expected rising price of fossil fuels
    - Rising interest in bio-based raw materials
    - Declining returns of forestry industry
    - High NOx emissions from Forestry
- Inputs/Resources:
    - Funding by CDI Programme
    - Other forms of support from Vinnova
    - Leadership of Rise Processum
    - Cross-sectoral network of scientists and technicians
    - Positive expectations towards bioeconomy
    - Previous collaboration in preceding projects (Vinnväxt)
- Activities:
    - Networking and knowledge exchange
        - Broadening and deepening network of chemical companies using purified methanol as an input
    - Identifying potentially beneficial value chains
    - Pre-Feed-studies and system analyses
    - R&D on pulp mills
    - Technical and economic evaluations
    - Pilot equipment testing
- Transformative outcomes:
    - Increased knowledge of the sustainability implications of paper production and use (first-order learning)
    - Networking through the participation in the project of a wide range of diverse stakeholders

- Expected impacts:
    - Reduced environmental impact through reduced NOx emissions, and reduction of greenhouse gases
    - Reduced emission-related costs for sulphate pulp use.

To build such a ToC using in an inclusive, participative way and to use it as a formative evaluation tool to enhance the probability of a programme to generate or contribute towards transformative change is much more complex than assessing the success of the project by measuring it against a limited set of technological performance indicators.[8] Although both approaches can be called "evaluations", there is little in common between them: they have different objectives, the communities involve are different as they are the processes and methods used. The formative approach we proposed can be seen as a tool for continuous policy design and implementation that attempts to preserve and pursue the systemic transformative goals that the policy is seeking.

## 5. Conclusions

We have argued that an evaluation approach for TIPs needs to be formative, aiming to improve the definition and implementation of the interventions under evaluation and involving the policy participants. This requires evaluation to be conducted in real-time, as a form of constructive monitoring. To be able to assess in real-time the degree to which the interventions are progressing towards the achievement of long-term systemic goals, the evaluation approach needs to be underpinned by both generic and specific 'Theories of Change' (ToCs). To enable reflexivity such ToCs need to be flexible, and should be revisited as part of the formative, real-time evaluation processes. The approach proposed includes a three-step procedure to build and revise the specific ToCs required to guide the evaluation process.

Although ToCs are common in policy evaluation in other domains (for instance in development), they were seldom used in the evaluation of innovation policies. Our interaction with STI policymakers suggested the importance to anchor evaluation on a generic ToC that would help build a common rationale and theory-base justification for TIPs: a stylised view of the transformative change processes derived from transitions theory (Grin et al., 2010, Markard et al., 2012). The resulting approach is innovative and provides an answer to the problem of assessing the downstream contributions and impact of current policy interventions, in a way that is coherent with TIPs principles.

In conclusion, we argue that, instead of acting as a perfunctory check at specified points of the project, TIP evaluation at its most effective should be a key element of the policy definition and implementation process, across different policy levels. As shown with the example of the Forest Chemistry project, the level of effort required for this type of evaluation is of a different order of magnitude from the evaluative analysis that supports the archetypical approach to summative evaluation. Yet, the role and function of these types of evaluation are very different: in our formative approach evaluation is part and parcel of a different way of defining and implementing policy, through which the

---

[8] Even for an exploratory project of the kind described here, the assignment involved the review of 9 scientific and technical reports produced by the initiative and a one-day workshop in January 2019 with the participation of 6 Vinnova officials and 4 research team members.

different stakeholders in a policy monitor and reassess policy results as they happen. It is a form of Real Time monitoring[9] embedded in the policy process.

Developing a new approach for transformative evaluation is a reflexive, participatory process that is interwoven at all stages of the policy process. As the policy evolves and adapts, so will the evaluation. Ultimately, TIP experiments will need a new evaluative strategy that must be co-created through the same actors who conduct the policy experiments.

## References

Abma, T. A. (2005) 'Responsive evaluation: Its meaning and special contribution to health promotion', *Evaluation and Program Planning,* 28, pp. 279-289.

Allmér, C. (2017) *Expect the Unexpected. Differences in Vision and Expectation in the Developing Bioeconomy.* Master's, Lund University [Online] Available at: http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=8917414&fileOId=8917415 (Accessed: 30 December 2019).

Arkesteijn, M., van Mierlo, B. and Leeuwis, C. (2015) 'The need for reflexive evaluation approaches in development cooperation', *Evaluation,* 21(1), pp. 99-115.

Arnold, E. (2004) 'Evaluating research and innovation policy: a systems world needs systems evaluations', *Research Evaluation,* 13(1), pp. 3-17.

Belcher, B. M., Davel, R. and Claus, R. (2020) 'A refined method for theory-based evaluation of the societal impacts of research', *MethodsX,* 7, pp. 100788.

Bos, B. and Grin, J. (2008) '"Doing" Reflexive Modernization in Pig Husbandry:The Hard Work of Changing the Course of a River', *Science, Technology, & Human Values,* 33(4), pp. 480-507.

Caffrey, L. and Munro, E. (2017) 'A systems approach to policy evaluation', *Evaluation,* 23(4), pp. 463-478.

Cohen, S. and Franco, R. (1992) *Rationalising social policy: evaluation and viability.* Santiago: CEPAL.

Connell, J. P. and Kubisch, A. C. (1998) *Applying a Theory of Change Approach to the Evaluation of Comprehensive Community Initiatives: Progress, Prospects and Problems*: Aspen Institute.

Daedlow, K., Podhora, A., Winkelmann, M., Kopfmüller, J., Walz, R. and Helming, K. (2016) 'Socially responsible research processes for sustainability transformation: an integrated assessment framework', *Current Opinion in Environmental Sustainability,* 23, pp. 1-11.

Daniels, C., Schot, J. W., Chataway, J., Ramirez, M., Steinmueller, E. and Kanger, L. (2020) 'Transformative Innovation Policy: Insights from Colombia, Finland, Norway, South Africa and Sweden', in Cele, M.B.G., Luescher, T.M. & Fadiji, A.W. (eds.) *Innovation policy at the intersection. Global debates and local experiences*: HSRC Press.

---

[9] Another example of real time "impact assessment" can be found in Joly, P.-B., Matt, M. and Robinson, D. K. R. (2019) 'Research Impact Assessment: from ex post to real-time assessment', *fteval Journal for Research and Technology Policy Evaluation*.

Earl, S., Carden, F. and Smutylo, T. (2001) *Outcome Mapping. Building Learning and Reflection into Development Programs.* Ottawa: International Development Research Centre.

Edler, J., Ebersberger, B. and Lo, V. (2008) 'Improving policy understanding by means of secondary analyses of policy evaluation', *Research Evaluation,* 17(3), pp. 175-186.

Eschenbach, R., Hafkesbrink, J. and Lütz (1995) 'Evaluation of hte Promotion of Pre-competitive Research and Development Cooperation in the Case of Microsystems Technology (MST)', in Becher, G. & Kuhlmann, S. (eds.) *Evaluation of technology policy processes in Germany*: Springer-Science-Business Media, B.V., pp. 165-196.

Fuenfschilling, L., Bauer, F. and Clemente, J. (2017) *Transformative Innvoation Learning History: Sweden. Challenge-driven Innovation: A new avenue for transformative policy at Vinnova*: Transformative Innovation Policy Consortium. Available at: https://www.tipconsortium.net/wp-content/uploads/2019/05/TIPC_Sweden_InnovationHistory_CDI_Final-4.pdf.

Geels, F. W. (2002) 'Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study', *Research Policy,* 31(8), pp. 1257-1274.

Geels, F. W. (2010) 'Ontologies, socio-technical transitions (to sustainability), and the multi-level perspective', *Research Policy,* 39, pp. 495-510.

Geels, F. W. and Schot, J. (2007) 'Typology of sociotechnical transitions pathways', *Research Policy,* 36, pp. 399-417.

Ghosh, B., Kivimaa, P., Ramirez, M., Schot, J. and Torrens, J. (Forthcoming) *Transformative Outcomes: Assessing and reorienting experimentation with transformative innovation policy*.

Giddens, A. (1984) *The Constitution of Society. Outline of the Theoru of Structuration.* Berkeley: University of California.

Grin, J., Rotmans, J. and Schot, J. (2010) 'Transitions to Sustainable Development: New Directions in the Study of Long Term Transformative Change', *Transitions to Sustainable Development: New Directions in the Study of Long Term Transformative Change*.

Heiskanen, E. and Matschoss, K. (2018) 'Evaluating Climate Governance Experiments', in Turnheim, B., Berkhout, F. & Kivimaa, P. (eds.) *Innovating Climate Governance: Moving Beyond Experiments*. Cambridge: Cambridge University Press, pp. 182-200.

Hildén, M., Jordan, A. and Rayner, T. (2014) 'Climate policy innovation: developing an evaluation perspective', *Environmental Politics,* 23(5), pp. 884-905.

Hobday, M. (1988) 'Evaluating collaborative R&D programmes in information technology: the case of the U.K. Alvey programme', *Technovation,* 8, pp. 271-298.

Holzer, J. M., Carmon, N. and Orenstein, D. E. (2018) 'A methodology for evaluating transdisciplinary research on coupled socio-ecological systems', *Ecological Indicators,* 85, pp. 808-819.

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

Joly, P.-B., Matt, M. and Robinson, D. K. R. (2019) 'Research Impact Assessment: from ex post to real-time assessment', *fteval Journal for Research and Technology Policy Evaluation*.

Kivimaa, P., Hildén, M., Huitema, D., Jordan, A. and Newig, J. (2017) 'Experiments in climate governance – A systematic review of research on energy and built environment transitions', *Journal of Cleaner Production,* 169, pp. 17-29.

Kivimaa, P. and Kern, F. (2016) 'Creative destruction or mere niche support? Innovation policy mixes for sustainability transitions', *research Policy,* 45(1), pp. 205-217.

Kuhlmann, S. (1999) 'Distributed intelligence: combining evaluation, foresight and technology assessment', *IPTS Report,* No. 40(Special Issue: Evaluation and Research Activities), pp. 16-22.

Luederitz, C., Schäpke, N., Wiek, A., Lang, D. J., Bergmann, M., Bos, J. J., Burch, S., Davies, A., Evans, J., König, A., Farrelly, M. A., Forrest, N., Frantzeskaki, N., Gibson, R. B., Kay, B., Loorbach, D., McCormick, K., Parodi, O., Rauschmayer, F., Schneidewind, U., Stauffacher, M., Stelzer, F., Trencher, G., Venjakob, J., Vergragt, P. J., von Wehrden, H. and Westley, F. R. (2017) 'Learning through evaluation – A tentative evaluative scheme for sustainability transition experiments', *Journal of Cleaner Production,* 169, pp. 61-76.

Magro, E. and Wilson, J. (2013) 'Complex Innovation Policy Systems: Towards an Evaluation Mix', *Research Policy,* 42, pp. 1647-1656.

Markard, J., Raven, R. and Truffer, B. (2012) 'Sustainability transitions: An emerging field of research and its prospects', *Research Policy,* 41(6), pp. 955-967.

Molas-Gallart, J. and Davies, A. (2006) 'Program theories and evaluation practice: the experience of Science, Technology and Innovation policies in Europe', *American Journal of Evaluation,* 27(1), pp. 64-82.

OECD (2016) *OECD Reviews of Innovation Policy: Sweden 2016*, Paris: OECD Publishing. Available at: https://www.tipconsortium.net/wp-content/uploads/2019/05/TIPC_Sweden_InnovationHistory_CDI_Final-4.pdf.

Popper, R., Velasco, G. and Popper, M. (2017) *CASI-F: Common Framework for the Assessment and Management of Sustainable Innovation, CASI project report. Deliverable 6.2*. Available at: http://www.casi2020.eu/library/deliverables/ (Accessed: 27 February 2018).

Rotmans, J., Kemp, R. and Van Asselt, M. (2001) 'more evolution than revolution: transition management in public policy', *Foresight - The journal of future studies, strategic thinking and policy,* 3(1), pp. 15-31.

Schot, J., Kivimaa, P. and Torrens, J. (2019) *Transforming Experimentation. Experimental Policy Engagement and their Transformative Outcomes*.

Schot, J. and Steinmueller, W. E. (2018) 'Three frames for innovation policy: R&D, systems of innovation and transformative change', *Research Policy,* 47(9), pp. 1554-1567.

Schot, J. W., Daniels, C., Torrens, J. and Bloomfield, G. 2017. Developing a Shared Understanding of Transformative Innovation Policy. *TIPC Research Brief.* Brighton.

Schot, J. W. and Kanger, L. (2018) 'Deep transitions: Emergence, acceleration, stabilization and directionality', *Research Policy,* 47(6), pp. 1045-1059.

Smallman, M. (2018) 'Citizen Science and Reponsible Research and Innovation', in Hecker, S., Haklay, M., Bowser, A., Makuch, Z., Vogel, J. & Bonn, A. (eds.) *Citizen Science: Innovation in Open Science, Society and Policy*. London: UCL Press.

Smith, A., Voß, J.-P. and Grin, J. (2010) 'Innovation studies and sustainability transitions: The allure of the multi-level perspective and its challenges', *Research Policy,* 39(4), pp. 435-448.

Smutylo, T. 'Crouching Impact, Hidden Attribution: Overcoming Threats to Learning in Development Programs', *Block Island Workshop on Across Portfolio Learning*, 22-24 May.

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

Taanman, M. (2014) *Looking for Transitions. Monitoring approach for sustainable transition programmes.* Doctor, Erasmus Universiteit Rotterdam, Rotterdam.

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

van Mierlo, B., Arkesteijn, M. and Leeuwis, C. (2010a) 'Enhancing the reflexivity of system innovation projects with system analyses', *American Journal of Evaluation,* 31(2), pp. 143-161.

van Mierlo, B., Regeer, B., Vam Amstel, M., Arkesteijn, M., Beekman, V., Bunders, J., de Cock Buning, T., Hoes, A.-C. and Leeuwis, C. (2010b) *Reflexive Monitoring in Action. A Guide for Monintoring System Innovation Projects.* Boxtel, the Netherlands: Boxpress.

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), pp. 1037-1047.

Weiss, C. H. (1998) *Evaluation. Methods for studying programmes and policies.* London: Prentice Hall.

Wilson-Grau, R. and Britt, H. (2013) *Outcome Harvesting*, Cairo: Ford Foundation Middle East and North Africa Office.