Evaluating Societal Effects of Transdisciplinary Co-production Processes: Final Report

Report developed for Mistra Urban Futures
By Steve Williams, PhD Candidate, University of British Columbia

Table of Contents

Evaluating Societal Effects of Transdisciplinary Co-production Processes: Final Report 1
Table of Contents 2
Introduction and project context 3
Literature review methods 3
Challenges in Evaluating Societal Impact at Mistra Urban Futures 4
Conceptions of co-production 5
Co-production in community-university partnerships 5
Evaluating co-production processes 6
Progress to Sustainable Urbanisation 11
Attribution 11
Evaluation Framework in Practice 12
Recommendations 15
Acknowledgments 19
References 20
Appendix A: Integrated Evaluation Framework 23

Introduction and project context

The vision and mission of Mistra Urban Futures clearly set out the goals of the programme:

> **Vision**: Sustainable urbanisation where cities are fair, green and accessible.
>
> **Mission**: To generate and use knowledge for transitions towards sustainable urban futures through reflective co-creation at local and global levels." (Mistra Urban Futures, 2015)

The programme has further defined these objectives of "working towards fair, green and inclusive cities "FAIR Cities: Securing urban equity, social inclusion, cultural diversity and urban commons. GREEN Cities: Managing resource constraints, urban environments, ecosystems and climate change sustainably. ACCESSIBLE Cities: Promoting effcient and equitable access to urban qualities, opportunities and services" (Mistra Urban Futures, 2015, p. 10) and defined a set of strategic objectives "I: Deliver evidence-based outcomes that address the challenges facing cities, and which make a difference in practice; II: Diversifying the Centre's research base and forging strategic partnerships with selected international organisations" (p. 13-14) and 5 knowledge themes:

1. Sustainable spatial urbanisation and urban qualities
2. Urban social sustainability
3. Integrating social and ecological systems
4. Sustainable urban lifestyles
5. The role of enterprise and civil society in sustainable urban development." (p. 17).

Most relevant to this report, the 2015 Mid-term Review report suggests four key objectives related to co-production: Applying co-creation in practice, rooting co-production in the LIPs, cross-LIP collaboration and evaluation around coherence of new themes and perspectives and finally contributing to a global agenda for sustainable cities (Reepalu et al. 2015, pp. 23-24).

While the Mistra Urban Futures programme has invested substantially in QME processes, the question of how to evaluate the societal effects of transdisciplinary co-production processes remains open. The Mid-term Evaluation conducted in 2015 (Reepalu et al. 2015, pp. 12-15) placed an emphasis on capturing societal impacts and proposed a preliminary evaluation framework building on Wiek et al. 2014. The evaluation approach used interviews in three of the centres (KLIP, GMLIP, and CTLIP) and document reviews to generate a frequency count of "each category of outcomes or impacts…in the narrative" (p. 13). The evaluation framework proposed in this report extends work by Wiek et al. and attempts to provide a set of qualitative and quantitative indicators of societal effect that can be more precise and wide-ranging than a frequency count of terms. In addition, this framework captures impact of co-production processes along with the cumulative impact of the actual outcomes of implemented results from research undertaken by the LIPs.

Literature review methods

For this report, I reviewed established literatures (Klein, 2008; Walter, Helgenberger, Wiek, & Scholz, 2007; Wiek, Talwar, O'Shea, & Robinson, 2014) and more recent literature (2016-2017) with a focus on evaluating co-production processes and literature on community-university collaborations. In addition, I reviewed previously published documents from Mistra Urban Futures

including the 2015 Mid-term Report, the 2016-2019 Strategic Plan and QME planning documents. Additional literatures include evaluation of the impacts of participatory processes and knowledge mobilization. Through this review, I collected indicators specific to co-production processes then integrated into, and extended, an existing societal evaluation framework (Williams 2017).

There has been much work already done on the evaluation of the co-production processes themselves (e.g. understandings of co-production, roles of participants) and recommendations for successful projects (Hansson & Polk, 2017, Cvitanovic et al. 2016, Fazey et al. 2014, Hegger et al. 2014 and Reyers et al. 2015). This literature points to the need for co-production process to be iterative, interactive and reflexive, provide transparent discourse and collaboration, and embed broad and diverse participation and engagement. This report is concerned specifically with societal impact (reviewed in Pajouhesh 2017) rather than process evaluation. However, as Walter et al. (2007) point out, there is a strong correlation between outcomes and "involvement as measured by the number of engagement activities that took place during the project". I recommend that evaluation of co-production process, design and implementation continue in parallel with societal impact assessment. Further research could integrate these two forms of evaluation to assess the relationship of Mistra programme design and process to outcomes. As Meagher et al. note, "by concentrating at least as much effort on understanding the processes of knowledge exchange involved in generating impact as much as the impacts per se it is possible to provide more immediate feedback to enhance delivery of knowledge exchange, increasing the likelihood that long-term impacts will be achieved" (2008; cited in Fazey et al. 2014, p. 218)

It is important to note the challenges of evaluating societal impacts. As Klein notes, "many long-term effects cannot be predicted or checked in five-year periods, let alone annual measures" (Klein, 2008). In addition, assessing societal impact through non-traditional measures can be difficult: "Even the idea of assessing the impact of research on decision-making is new within many academic disciplines, where reward structures rely primarily on the number of peer-reviewed publications (Bell et al. 2011; Roux et al. 2010)." (cited in Wall et al. 2017)

Challenges in Evaluating Societal Impact at Mistra Urban Futures

In addition to the challenges of evaluating societal impact in general, the Mistra Urban Futures programme poses some additional unique challenges. First, there is an inherent tension between the needs and goals of individual LIPs and the collective group that is MUF. Each LIP is a unique context and has its own approaches to achieving sustainable urbanization. At the same time, there is value in the collective work where a group of LIPs can more effectively highlight issues at a global level. Evaluation needs to be happening at two different levels - within the LIPs and the Centre as a whole where the whole is more than the sum of the parts. For example, by articulating outcomes through the Centre, the secretariat is able to "act up" into international and global agendas, drawing on comparative work and the work of the LIPs. The Centre is not duplicating the work of the LIPs but as a partner can organize and support joint events drawing on the work of the other LIPs to share good practice and ensure that LIPs gain legitimacy and recognition from participation in MUF. Second, there is the perennial challenge of delivering long-term impact when working with short-term budget cycles. Many of the outcomes and effects of MUF work take time to come to be realized. Policy changes, organizational changes, shifts in public behaviour and attitudes can change over years or even decades. This poses a challenge when projects are working on short-term funding and funders are asking for immediate impact measures.

Conceptions of co-production

There has been much literature on co-production and co-creation of research and action in recent years (Jasanoff, 2004; Lemos and Morehouse, 2005; Pohl, et al, 2010). In this conception of knowledge development, researchers work with colleagues in the private, public, and/or civil society sectors in the articulation of research questions, in the undertaking of the research itself, and interpretation and use of the results (Robinson and Tansey, 2006; Talwar et al, 2011). The co-production process purports to have many benefits and "Previous research has shown that taking a collaborative approach to knowledge development is more likely to result in science that is used by decision-makers" (Wall, Meadow, & Horganic, 2017). This model aligns directly with the concept of transdisiplinarity, defined by Polk as "different types of knowledge production for societal change based on in-depth collaborative processes that integrate knowledge from different disciplines (interdisciplinary) with values, knowledge, know-how and expertise from non-scientific sources. The focus on societal change is a critical one. Mistra Urban Futures is very much taking a strong co-production approach which implies a need to capture not just knowledge produced but the visions of a future world that are being co-produced. In this respect, co-production while important, is not the sole end goal but also means to achieving the MUF vision.

Bremer & Meisch (2017) describe a set of lenses of co-production. The lens chosen will impact evaluation measures and conceptions of what "good" co-production means. These are not better than one another but different viewpoints that may even be different between LIPs. We need to be explicit about our approaches and how that shows up in evaluation measures – e.g. broad bucket for co-production with a range of possible indicators. Based on my analysis, MUF combines two normative approaches to co-production: Iterative interaction and Social Learning. The former emphasizes knowledge translation and use, while the second foregrounds the importance of co-learning between participants (both individuals and organizations). An evaluation framework should include indicators that allow for an assessment of these approaches to co-production.

Finally, it is important to note that the capacity for co-production is not innate but can be learned (Hegger et al. 2012). Literature suggests a set of capacities for co-production. In addition to conditions such as diversity of actors and availability of resources (Hegger et al. 2012), capacities such as relationship quality, and social learning to support adaptive management approaches are also required (Reyers 2015). These capacities include researcher (e.g. social learning and stakeholder collaboration), financial (including funding plans beyond the initial project) and institutional (e.g. training and support for co-production initiatives) (Clark et al. 2016; Cvitanovic et al. 2016). Process evaluation should include measures that can help assess not only the capacity of a project team and its partners but also the capacity development programs that have (or have not) been put in place.

Co-production in community-university partnerships

Within the co-production literature, there is a sub-set that specifically focuses on the community-university relationship and highlighting the "importance of recognizing that both scholars and practitioners are, in fact, researchers." (Buick, Blackman, O'Flynn, O'Donnell, & West, 2016). However, the focus of each group may be different – what is often referred to as the 'research-

practice gap'. For example, "Scholars tend to be more concerned with rigor, and this drives their focus on the construction or refinement of theories or new knowledge underpinned by rigorous investigation, data collection, and advanced analysis methods, often in the quest to establish general principles. Practitioners tend to be more concerned with relevance—that is, the need for context and problem-specific research that aligns with real-life problems (Buick et al. 2016, p. 36). Buick et al. also propose a set of strategies for managing this tension: developing a clear shared purpose, enabling governance arrangements, joint approach to data analysis, joint approach to research dissemination, composition and management of the project team "including a range of views, disciplines, research methods, skills and levels of experience", building relationships, respect and trust. (2016, pp. 42-44). Again, these procedural and design elements are ones that can be assessed.

Evaluating co-production processes

Procedural vs. outcome evaluation (process vs. impact)

Evaluation frameworks tend to focus on the design and procedural elements of the process itself or on the outcomes of the process. Procedural evaluation looks at whether processes are inclusive, fair, and present unbiased information (Abelson et al., 2003; Black, Burkhalter, Gastil, & Stromer-Galley, 2008; Carpini, Cook, & Jacobs, 2004). Following Ableson et al. (2003), Rowe and Frewer (2005) propose a framework of fairness and competence/efficiency to measure the effectiveness of public engagement processes. Fairness is a measure of perception - do participants (and the wider public) believe the exercise was designed and conducted in a fair and representative manner – and so largely procedural. However, Rowe and Frewer extend the definition of competence to include how efficient the information flow is and how well that information is processed (i.e. into policy or other outcomes and objectives) – and so largely substantive. While measures of process effectiveness are quite well established, outcome measures are much less developed and will be the focus on my framework. Outcome evaluation looks at both impact of the process on participants (e.g. in terms of increased knowledge or level of civic participation), on the products of the process (e.g. reports) and the use of those products by decision makers (Barrett, Wyman, & Coehlo, 2012; Caddy, 2005; Gaventa & Barrett, 2010; Mutz, 2008; Pincock, 2012; Ryfe, 2005). A commonality of this literatures is that it comes from a governance perspective asking how citizen engagement will link to policy and decision making (Darcy Riddell, Personal Communication, April 12, 2017). While policy effects and decision quality are important, there are many other possible dimensions of societal impact.

Evaluating Societal Effects

Literature on evaluating societal impact (Wiek et al. 2014, Luederitz et al. 2016, Cornish 2013, etc.) suggests that we evaluate first-order effects such as enhanced capacity, network and usable products (e.g. action plans, web sites, new technologies), second order effects such as structural changes (e.g. new policies, organizational changes), decisions and actions, the development of new visions and imaginaries and changes in practice (Williams 2017).

How then might one define and measure the societal impact of deliberative fora? Wiek et al. (2014), Cornish (2013), Robinson (2006) and others have extended the literature on the impact of participatory research (Baldwin, 2000; Blackstock, Kelly, & Horsey, 2007; Currie et al., 2005; Kasemir, Jager, Jaegar, & Gardner, 2003; Robinson, Burch, Talwar, O'Shea, & Walsh, 2011; Talwar, Wiek, & Robinson, 2011; Walter et al., 2007) to propose a new evaluative framework.

This framework categorizes societal effects into first-order (the short term "splash" from a specific event or process) such as enhanced capacity, network and usable products (e.g. action plans, web sites, new technologies) followed by second-order ("the ripples" which are bigger impacts that typically take longer to appear) such as structural changes (e.g. new policies, organizational changes), decisions and actions. This approach broadened the categories of effect from previous work in the field by including structural changes as evaluative criteria. Wiek et al.'s framework also acknowledges the challenges of attributing effects due to time delays between processes of deliberation or equivalent events that have occurred and their effect. While the 1st and 2nd order effects framing is very useful, the definitions and indicators of 2nd order structural effects are not precise, for example conflating societal shifts such as norms and behaviour change with policy and institutional effects. These are very different types of effects and analysis would have more clarity by separating these effects.

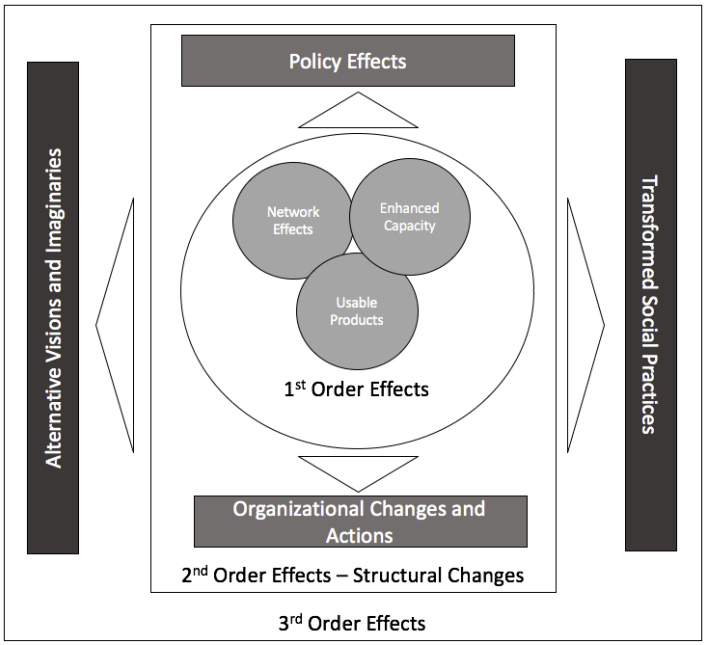
Wall et al. (2017) provide detailed indicators usefully clustered into inputs, process, outputs, outcomes and impacts. This allows us to more precisely define the type of policy impact that a project may have had. However, the focus of their impact measures is on the instrumental use by decision makers of knowledge produced in the project. Walter et al. (2007) also foreground this goal. This is an important factor to consider, but societal impacts include much more than this. For example, it is important to capture are "the relationships that allow co-production to happen and the new forms of knowledge, values, and social relations that emerge out of co-productive processes" (Filipe, Renedo, & Marston, 2017) and further "not only take the form of interactions between individuals and services, but also involve interactions between different rationales for participation and policy agendas, between different modes of knowledge production (e.g., knowledge based on biomedical evidence, clinical practice, or experience of illness), and between different kinds of value (e.g., economic value and values of equity and social justice)". However, Klein et al. (2008) point to the variability of goals, criteria and indicators as challenges of evaluating transdisciplinary projects. In light of this, they propose – in addition to specific impact measures – evaluating a project's responsiveness to variation and including long-term unpredictable impacts in evaluation frameworks.

Fazey et al. (2014) have a specific focus on the knowledge exchange dimension of co-production processes. Through a literature review, the authors propose indicators clustered into process, understanding, practice/policy change and impact of practice/policy change. Of these, the latter is the most interesting and also the most difficult to capture in the short/medium term. Leuderitz et al. (2016) also attempt to integrate these high level goals in framing sustainability transition experiments.

What might be a useful extension to the Wiek et al. model? Kearnes and Chilvers (2016) introduce the idea of participatory processes being "multiply productive" in that: "the productivity of participation goes way beyond discursive or linguistic outcomes to include material commitments (e.g., in bringing forward new technological commitments through 'grassroots' and distributed innovation in community energy projects, hackerspaces, design collectives, etc.), alternative visions and imaginaries (e.g., through forms of activism and artistic engagement), the potential for transformed social practices beyond the setting of specific participatory experiments (e.g., in relation to pro-environmental behaviour change initiatives), and so on" (2016, p. 40). Material commitments effectively include Wiek et al.'s 1st and 2nd order effects. The concepts of alternative

visions and potential for transformed social practices provides a set of effects that can be thought of as a 3rd order– how societies conceive of themselves and behave which can then include effects such as norm and behaviour change. Luederitz et al. (2015) have recently extended this concept further specifically attempting to measure impacts of sustainability transition experiments by asking whether the experiment strengthened socio-ecological integrity, enhanced livelihood sufficiency and opportunity and other measures of sustainability. However, this framing is evaluating specific interventions not the aggregate portfolio of experiments that are part of a transdisciplinary programme operating in multiple locations such as Mistra Urban Futures. This leads the framework to under-represent the importance of mutual reinforcement dynamics between experiments (Grin, 2011; Riddell, 2015) and where the programme fits into (i.e. both impacting and being impacted by) a broader set – or ecology – of processes (Chilvers & Kearnes, 2016). In addition, this framework does not address how a given process is contributing to a sustainability *transition* which is different from evaluating specific indicators of sustainability.

Following Wiek et al. (2014), **first order effects** are those most closely connected to a particular process and typically emerge in close temporal proximity to a given event or process. These effects are clustered into *usable products* (e.g. technologies or action plans), *enhanced capacity* (e.g. learning and co-production of knowledge) and *network effects* (e.g. new or strengthened networks) that are developed through the process.



Second order effects are, akin to outcomes, impacts on the system within which a given process is operating. *Economic*, *policy* (e.g new policies or changes in policy discourse) and *organizational* (e.g. new business models or changes in organizational strategy) effects may occur. These categories of effect tend to be further removed temporally from the initial experiment and direct causal attribution is more difficult.

Third order effects are those that transcend particular policy institutions or organizations and take place at a societal level. *Alternative visions and imaginaries* represent a different future, one that is different to the current prevailing view. This can take the form of new narratives that emerge from a process, development of shared visions or challenging existing conceptions of the social imaginary. *Transformed social practice* may occur as the result of, for example, pro-environmental behaviour activities of a process itself. Alternatively, these transformations may be supported/enabled by changes in the social imaginary as when changing conceptions of cleanliness led to radical changes in personal hygiene, bathing, clothes washing, etc. This is akin to definitions of social innovation which focuses on processes that "challenges and, over time, changes, the defining routines, resource and authority flows, or beliefs of the broader social system in which it is introduced".

Although presented here as three orders which implies a temporal flow (i.e. that second order effects are preceded by first order and third order effects follow from second order), societal effects are not as neatly defined as this. There are mutually reinforcing loops of influence wherein, for example, innovative products showcased as exemplars and demonstrations may help create shared visions of the future that can make policy change more likely/acceptable. In addition, categories of effect may be seen at different levels of analysis. For example, individual participants in a process may build capacity while at the same time contributing to capacity development within their organization. Similarly, changes in policy or new policy discourses may be achieved through individual participant action or transition experiment initiatives. The process itself may experience changes in structure and adaptation to changing landscape contexts. Finally, case-specific outcome measures may change over the life of the project (Fiester, 2011).

Co-production specific indicators

The literature provides a set of co-production specific indicators (see Appendix A). Note that each of these impacts may be assessed quantitatively (e.g. did the impact occur, what was the frequency) as well as the quality (e.g. what was the level of co-production, was there a lasting impact) (Fazey et al. 2014; Reepalu et al.. 2015). The impacts proposed fit well into the categories developed in previous work (Williams 2017). In this section I focus on co-production specific indicators. In the following section, I propose an integrated framework connecting co-production to participatory process and community-university sustainability evaluation literatures.

First order effects may be *products* such as written reports, research and practice overviews, (spreading knowledge) meetings with decision makers, workshops, seminars; funding applications; peer reviewed articles; technical reports/gray literature; media reports; creation of innovation and new ideas. *Capacity* impacts may include changes in ways of thinking, learning and working (distinguished between individual capacity building and mutual learning, and types of knowledge (system, goal, transformation); broadening of perspectives and knowledge, new understandings of subject area (e.g. practices of urban planning) transformation of knowledge with policy makers; changes in understanding (increased knowledge, change in attitudes, changes in thinking, new skills, increased confidence. Fazey et al. (2014) imply the need for a potential new category of knowledge exchange/distribution – e.g. provision of information to capture the flow of information within and outside the project. *Network* effects include new ways of working across disciplines and sectors; better understandings of roles of partner orgs; better understanding of how to make use of expertise and networks; collaborative funding applications; network building, community identification; national, international and local partnerships.

At the **second order** of impact, *policy* impacts may include effects on discourse leading to new perspectives and questions on the agenda and information use in decision-making". Note this information use may be conceptual (decision makers are better informed), justification (used to justify a predetermined decision), instrumental (to inform a new decision). Additional *policy* impacts can include distribution of knowledge; informing management decisions, policy actions or adaptation decisions; new evidence introduced into policy/strategy, changes in policy or practice leading to a further dimension of evaluating *impacts* of changes in policy or practice – e.g. are there notable improvements in ecological health, social and economic welfare, social equity, business performance, etc. *Organizational* effects include impact on education (e.g. project results

used in teaching materials); change in decision making processes. **Third order** effects may see changes in policy or *practice* and new shared *visions* of a possible future aligned with (or different from) the MUF vision.

Multi-site programme evaluation

A particular strength of the Mistra Urban Futures' learning network, consisting of LIPs and partners, is the potential for North-South mutual learning across different contexts. This also provides the ideal platform for broader communication of our findings, tools and methods for promoting transitions towards urban sustainability elsewhere – "our unique selling point and an essential requirement for fulfilling our mission." (Mistra Urban Futures 2015 p. 24). A paper led by Mistra Urban Futures articulates some of the challenges of developing indicators that are important, relevant and easy to report on (Simon et al., 2016). Each LIP will need flexibility to interpret this evaluation framework locally while maintaining fidelity to the overall approach.

An Integrated Framework

Appendix A shows a proposed integrated framework linking co-production specific indicators to indicators of societal effect. Table 1 below provides a summary view of the table. Each type of effect may have a range of indicators.

Table 1 Summary of Integrated Evaluation Framework

Orders of Effect	**Categories of Effect**		
1	*Usable Products* - Technologies and social innovations - Publications - Distribution of knowledge	*Enhanced Individual Capacity* - Acquired knowledge (individual or collective; systems/process) - Personal change - Decision making capacity	*Network Effects* - Networks created or expanded - Community trust created or expanded - Community identity
2	*Policy Effects* - Policies/decisions made - New evidence and actors included in policy decisions - Solutions implemented	*Organizational Changes and Action* - Changed context for new and ongoing work - New organizations and business models - Change in decision making processes	
3	*Alternative Visions and Imaginaries* - Shifts in public narrative - Collective purpose and vision - Greater social cohesion across groups	*Transformed Social Practices* - Norm change and/or adoption - Inclusion of new actors and issues in public spaces and discourse - New space for innovation and experimentation	

Using this framework provides flexibility in approaching evaluation. For example, the Manchester LIP may choose to focus on indicators of *Enhanced Capacity* and *Policy Effects* while the Cape Town LIP may focus on *Network Effects* and *Organizational Changes*. The choice of which categories of effect to focus on should be driven by the organizational and project goals for each LIP and the Centre. For example, the Centre has a mandate to develop international partnerships,

facilitate knowledge sharing across LIPs, and link the work of the LIPs to global conversations on sustainable cities. In this case, focusing on *Network Effects* and *Enhanced Capacity* would be appropriate. In addition, capturing the impact of the work of the LIPs within Mistra Urban Futures on contributing *Alternative Visions and Imaginaries* of sustainable and just cities would be appropriate.

Progress to Sustainable Urbanisation

In parallel with societal effects evaluation, it is important consider whether these effects are in service of Mistra Urban Future's vision of sustainable urbanization. This transition is fundamentally long-term which needs long term thinking and commitment when funding and electoral cycles are multi year. This framework provides a methodology to look for early indications of systems change such as new policies implemented. However, having a new policy in place is one thing but evaluating the effect of that policy and whether that policy is leading towards more sustainable and just cities is something else. Might there be general indicators such as total GHG emissions, diversity, income inequality or social indicators that can serve to help guide us?

The UN Sustainable Development Goals (SDGs) may provide a way to bridge this gap. This set of goals covers environmental, social and economic indicators. The goals are widely known and used in organizations around the world. I recommend that further work be conducted that correlates societal effect measures with the SDGs. This would provide a connection between the measures of societal effect and progress towards MUFs vision and goals. For example, targets within the SDG 11 "Make cities and human settlements inclusive, safe, resilient and sustainable" would align with the Mistra Urban Futures vision. Sample targets include:

- **11.2 -** By 2030, provide access to safe, affordable, accessible and sustainable transport systems for all improving road safety, notably by expanding public transport, with special attention to the needs of those in vulnerable situations, women, children, persons with disabilities and older persons
- **11.B -** By 2020, substantially increase the number of cities and human settlements adopting and implementing integrated policies and plans towards inclusion, resource efficiency, mitigation and adaptation to climate change, resilience to disasters, and develop and implement, in line with the Sendai Framework for Disaster Risk Reduction 2015-2030, holistic disaster risk management at all levels
- **11.C -** Support least developed countries, including through financial and technical assistance, in building sustainable and resilient buildings utilizing local materials[1]

More work will be needed to thoroughly map the complete set of societal effect indicators to the SDG target and indicator framework but this is beyond the scope of this report.

Attribution

A final note on attribution of effects to a specific cause. As we know, societal outcomes are highly overdetermined by many factors (to greater or lesser extent per LIP). Nevertheless the actions of LIPs can contribute to these outcomes. How can we assess how they have made a contribution to a large scale outcome such as carbon emissions dropping. Data analysis methods such as Process

[1] https://sustainabledevelopment.un.org/sdg11

Tracing and Contribution Analysis can be helpful in assessing the contribution of a given process to decision making processes. However, linking that decision (e.g. a new public transportation policy) to eventual outcomes (e.g. carbon emissions reduction) is a greater challenge. Additional research is required in this area and will form a large part of my PhD Thesis work.

Evaluation Framework in Practice

The following section (excerpted from Williams 2017) illustrates how this framework has been applied in the context of the Alberta Energy Futures Lab (EFL). The EFL consists of 35 Alberta energy system leaders who are participating in a collaborative 5-year leadership development and rapid prototyping program designed to answer the question "How can Alberta's leadership position in today's energy system serve as a platform for the energy system the future requires of us?" This illustration serves as an example of how the framework might be implemented within Mistra Urban Futures and the types of effects that can be captured.

Capturing societal effects in the Alberta Energy Futures Lab (EFL)

Data Collection

Prior to the start of the sustainability transition experiment process, 38 baseline interviews were conducted in Summer of 2015. This interview consisted of questions asking what prospective participants see as the biggest barriers to and opportunities for transition, what they see as their role in the transition and what success (and failure!) would look like for the process. This data has been used as a baseline to track shifts in participant attitude and learning and to help with design of the lab process. During the process, semi-structured interviews are conducted with all participants by phone or in person (if possible) twice per year. These track individual, prototype, organization and network changes/impacts (first order effects) and to understand participants' own insights into emerging or potential societal shifts (second order effects) they regard as having occurred due to efforts of the lab process and/or other possible explanations. Interviews focus on impacts, and the mechanisms through which the impacts occurred (e.g. actions taken by participants), primarily at the levels of the individual participant, their organization, the prototype team(s) they are working with, and social and professional networks of which they are a part. Unstructured questions include asking participants to identify particular powerful actions that people think have the most impact and results that might deserve further investigation. For example, asking what has been the most surprising or unexpected result of the EFL so far and where participants are seeing signs the energy transition in Alberta is starting to happen. In addition, participants are asked for their own characterization of the EFL – what do they see as the goal of the process. These questions help identify areas of impact at the structural and cultural levels that can be explored further using other research methods as described below. Structured questions include asking whether participants have personally experienced new relationships, connections and collaborations with Lab Fellows, new understandings of the energy system in Alberta or shifts in how they think about their own role or contribution to the energy transition. These structured questions are designed to elicit data from participants on actions or changes at different levels of analysis. Other questions ask about new actions that have been taken or new insights generated at the prototype (working group), organization or professional/social network levels.

Early Indicators of Impact

The Energy Futures Lab process is still ongoing so it is too early to talk about all societal impact but first rounds of participant interviews have surfaced insights and show early indications of societal impact.

First order effects

First order effects take the form of usable products, enhanced capacity and network effects. There is strong evidence of impact in all three of these categories.

Usable products

There have been a series of products from the EFL including an energy system vision document, a series of blog posts, newspaper op-eds and public events. Changes in expectations for and evaluation of prototypes. Participants brought different innovative technologies to the EFL and some new ones have emerged. In particular, a set of technology innovations that have potential to radically reduce the GHG emissions from oilsands production have been proposed as part of a prototype called AOSTRA 2.0. The current portfolio of initiatives does reflect – to varying degrees – a focus on transition and EFL criteria for prototype selection. However, the mutually reinforcing potential of initiatives is unclear. There is also a lack of agreement on the incremental vs. disruptive nature of the innovations proposed.

Capacity

A majority of participants reported new insights about themselves along with new insights about the energy system and transition. We have seen some increased insight and empathy into the perspectives, narratives of other actors in the system (e.g. "I did not know that energy companies were so interested in sustainability"). Approximately 50% of participants feel a renewed sense of hope due to EFL and political events around transition (mostly ENGOs and Associations) while Approximately 25% report an increase in confidence/boldness in their role and in discussing energy transition in public. We have also seen adoption of process techniques such as backcasting within participant organizations.

Network effects

There are many examples of cross-connections between Fellows and we see about 65% connecting on non-prototype activities such as information sharing, education/training, joint projects and brokering relationships. Connections include discussing new market opportunities, making board recommendations, and partnering on projects. Note that these connections are separate from collaborations that are the happening with formalized prototype or working groups within the EFL.

Second order effects

Second order effects take form of policy/economic and organizational effects. There is strong potential for the EFL to impact policy but not much direct evidence of this happening. This is to be expected given the time scales of policy change and implementation. There is stronger evidence of the effect that the EFL is having on shifting organizational structures.

Policy effects

While it is likely too early to expect to see direct policy influence as a result of the EFL, there are many potential avenues where this effect could take place. For example, two EFL Fellows are part of the Government of Alberta's Energy Efficiency Panel and two Steering Committee members are on the Climate Technology Task Force. At a municipal level, there are a number of Fellows sitting on the City of Edmonton's Energy Transition Advisory Panel. These are all areas where the EFL could have substantial policy influence. I will continue to track this influence. Another potential policy influence is the prototype "360 Policy Lab" which will provide strategic policy advice on energy issues in Alberta to government and industry using the expertise of Fellows and process techniques of the EFL. This has the potential to influence both policy discourse along with specific policy decisions.

Organizational effects

There have been a number of examples of participant organizations starting to adopt process and substantive innovations/ideas from the EFL. For example, the Calgary Chamber of Commerce has adopted EFL engagement processes and changed the way they deliver public events making them much more interactive and participatory. Another example has been the formal adoption of the AOSTRA 2.0 prototype by the Canadian Oilsands Innovation Alliance (COSIA). COSIA is a collaborative of the major oilsands companies that jointly share intellectual property. Initially resistant to the idea of championing a set of technologies and ideas that emerged from outside of the COSIA process and had an explicit sustainability focus, COSIA changed its definition of innovation and (at the request of EFL Fellows working on the project) took ownership of the AOSTRA 2.0 prototype. Not only will this give resources and industry credibility to the ideas emerging from the EFL, it represents an example of a shift in operating model and institutional framework in an organization at the centre of Alberta's oil industry. Continuing series of workshops inside organizations such as Suncor and the National Energy Board hold promise for shifting organizational culture and decision making. Data collection on these activities is still ongoing.

Third order effects

Third order effects including alternative visions and imaginaries along with transformed social practice are, like second order effects, more difficult to assess in the short term. However, we are seeing early indications that the EFL may lead to these types of effect and are seeing some effects at the participant level.

Alternative visions and imaginaries

Shifts in public narrative has always been one of the EFL's explicit aims and the EFL team's approach to this has evolved over time. "EFL in the public" encapsulates this approach. The EFL Showcase event in Calgary that opened this paper is an example of the kind of initiative that will receive more attention in the next phase of the EFL. Rather than launch public education campaigns, the intent of the team is to seed, highlight, and showcase new narratives in the public discourse on energy in the Province. This is a highly polarized debate so just having the diversity of viewpoints in one group and on the same stage is an illustration of cohesion across groups that are typically seen on opposite sides of the debate. Public documents such as the Vision and Pathways that was agreed to by all of the Fellows is another illustration of an attempt to create a

new shared vision of the future of energy in the Province. It remains to be seen how widely adopted these new narratives and visions may be however the attendance of 450 people at the first major EFL public event may be an early indicator that the messages of the EFL are resonating with a wider public.

Transformed social practice

We can see norm and behaviour change within the space of the EFL itself. Participants have learned techniques to have difficult conversations across right/left, industry/NGO and other boundaries. Indeed these conversations have led to deepened personal connections among the Fellows. However, this may or may not translate into wider norm shifts in the province. Despite that, the EFL is a new space for innovation and experimentation that, once again, has the potential to shift norms and has already increased the capacity within EFL participants for collective action.

Recommendations

In this section I provide general recommendations on implementing a societal effects evaluation framework at MUF along with specific recommendations on data collection and analysis methodologies (also excerpted from Williams 2017).

General Recommendations

1. Use this common framework across the MUF Centre and LIPs while allowing each LIP to concentrate on measures that are most relevant to their projects. This approach simplifies the work required by each LIP, allows for an aggregated view of impacts across the LIPs, and provides data for comparison of effects between LIPs.
2. Each LIP and the Centre should prioritize indicators based on their unique goals and local context. This will generate a hierarchy of indicators, so that there could be a spectrum from a more high-level evaluation that is simpler and cheaper to undertake, to a full-bore complete evaluation that can be done if resources are available.
3. Link the societal effects framework with the SDGs and integrate with ongoing MUF work on realising just cities and defining what just cities entail. This integration will allow MUF to understand and articulate not just the effects of the programme but how those effects are contributing to the overall goal of sustainable urbanization. This would also provide an overall framework for interpreting the data collected and provide an overall assessment of impact and outcomes from MUF. MUF has already completed work mapping various projects to SDGs. This work could be extended to the level of indicators.
4. As much as possible, integrate societal effects evaluation with existing and ongoing QME and process evaluation work. While there are specific methodologies and indicators for societal effect, there is opportunity to leverage existing staff, skillsets and processes for evaluation. Such an integration will provide a full picture of the impact of MUF.

Mobilizing Knowledge

When collecting, interpreting and analyzing the data collected, thought should be given to how and where that data may be used. Potential audiences and dissemination formats for this data are shown in the table below. Note that these publications, dissemination of knowledge and influence on audience understanding and decision making are themselves effects of the programme that can, and should, be captured in the evaluation framework.

Audience	Use	Formats
LIPs	Understanding local effects Learning from other LIPs Demonstrating impact	Reports Webinars Summary presentations
MUF Funders	Demonstrating impact of current projects Generating interest and funding for new LIPs and projects	Reports Webinars Summary presentations
Researchers	Sharing learning on value of MUF approach Demonstrating impact of MUF projects Furthering the field of impact measurement of complex projects	Journal articles Conference presentations Convened symposia
Policy Makers	Demonstrating value of MUF approach (i.e. transdisciplinary co-production)	White papers Conference presentations Policy maker workshops

Participatory Evaluation

Use evaluation process as way to build relationships and understanding among programme participants, across LIPs and between LIPs and Centre through co-designed evaluations. "Carefully designed and managed participatory evaluations become powerful tools for enhancing knowledge exchange, encouraging people to wrestle with epistemological differences and their assumptions. Such evaluations are useful for interdisciplinary endeavours that genuinely aim to build social relations." (Fazey et al. 2014, p. 218). Participatory evaluation can serve to facilitate this dialogue process and ensure that local cultural, research and project contexts are integrated into the evaluation process.

Data collection methodologies

A range of data collection methodologies can be used for data collection. Recommended is the semi-structured interview format described above for programme participants, programme design team (at Centre and within LIPs) and key bellwethers. In addition, document review (e.g. meeting minutes and reports) and media analysis can be useful in identifying societal effects. To generate more in-depth case studies, outcome harvesting (Koch, 2014; Quinlan, Kane, & Trochim, 2008; The World Bank, 2014) can be used. Outcome harvests have been used extensively by the World Bank in project evaluation and represent a broad approach to evaluating complex change initiatives. The focus of the method is developing "change stories" which can include one more indicators of societal effect and movement towards sustainable urbanisation. Interpretation of these stories is verified with participants who reported them and then the reported change or outcome is validated with an external participant. This helps address the significant difficulty of establishing a baseline for a complex change initiative where outcomes are diverse (i.e. trying to develop a baseline for every possible outcome in an efficient manner is practically impossible). Outcome harvesting can use multiple other evaluation methodologies such as ripple effect mapping, bellwether interviews, most significant change, or contribution analysis.

Data Analysis Methodologies

Interviews with participants are a very useful set of data but impacts are self-reported. Bellwethers are also useful but are often limited in their viewpoints. In addition, these are purely descriptive of changes and do not really get at whether the change happened as a result of lab process or would have happened anyway. It can be very difficult to create direct attributions between processes like these and societal impacts. I will attempt to get around this by looking for direct references from external sources like government reports, speeches, etc. and being very precise about tracing actors, observing where they exert influence and how. For example, a mention of the contribution of the MUF to policy development in a public speech by the a government official would be a direct reference. However, more work would be needed to ascertain how much impact the MUF actually had as opposed to the government wanting to bolster the image of themselves as a convening partner in a successful engagement process. The following methods represent an attempt to provide a causal link between reported and observed effects and the MUF itself.

There is a substantial amount of literature from the international development sector on project evaluation that attempts to capture long term impacts on policy, social welfare, etc. These literatures approach evaluation "inside out" - how project participants report impact or most significant change based on their own observations (White & Phillips, 2012; Davies and Dart, 2005) and "outside in" - how impacts are observed by evaluators that are explicitly outside of the process (George & Bennett, 2005; Reilly, 2010; Hughes and Hutchings, 2011). Both are useful to identify *potential* effects and triangulating self-reported outcomes. Two key methods are useful to get at 'causality' or the connection of impacts to mechanisms within the MUF. The first, process tracing, "explicitly sets out to discover the causes of observed effects" (White & Phillips, 2012) while the second, most significant change, privileges participant experience of change.

Process tracing

Process tracing (George and Bennett, 2005; Reilly, 2010) is used in political science research along by Oxfam (Hughes and Hutchings, 2011) and other international development organizations to validate causal hypotheses through collection of evidence for and against. One can propose a number of competing hypotheses about how the process led to a particular outcome. For example, that it was the convening of partners across ideological divides that led to innovative policy implementation or that it was actually the downturn in oil price that led to economic pressures making policy innovation more likely. Evidence is then collected (for example in the form of meeting minutes, policy documents, government briefings) in order to "overturn or substantiate rival hypothetical explanations" (White & Phillips, 2012).

Most significant change

Most significant change (MSC) methodology (Davies and Dart, 2005) "involves the systematic collection and selection of a purposive sample of significant-change stories. The stories themselves are elicited from programme participants by asking them to relate what significant changes (positive or negative) have occurred in their lives in the recent past [or in the case of the EFL, changes in their professional life], and enquiring why they think that these changes occurred and why they regard them as being significant" (White & Phillips, 2012). This technique takes an "inside out" approach and relies on participant reporting to define what impacts and causal chains took place. In this case, interviews with participants and bellwethers will surface what *they* think are the most significant changes as a result of the MUF and how this occurred. The combination

of both techniques, along with bellwether interviews and environmental scans, allows the evaluator validate reported and hypothesized causal links. However, it must be stated that the best result might be an assessment of contribution to change rather than causal attribution.

Contribution analysis

Contribution analysis is a method that helps test a theory of change, determine whether expected outcomes were achieve and evaluate whether a given process contributed those outcomes (Mayne, 2001; Mayne, 2011). By listing the assumptions behind the theory of change and gathering evidence to support (or disprove) those assumptions, the researcher can make claims as to the robustness of contribution analysis. This might be minimal – in effect stating that expected outputs were delivered, direct – making claims as to direct contribution of the process to outcomes, or indirect – making supported claims about indirect contributions to outcomes.

Context sensing and monitoring

In many ways, future studies are entirely about the present (John Robinson, Personal communication, September 2016). For example, the changing landscape of sustainable cities in a global context will have, and has already had, significant impact on the MUF and LIP design, process. It is therefore critical to track these context shifts through methods such as content analysis of a range of publically available media, government and industry reports. The aim here is to produce a history of narratives of sustainable urbanisation from different sectors and evidence of shifts in narratives or reported societal impacts within the broader context of sustainable cities at the LIP level and within a global context. This method is particularly useful in identifying changes in public narratives and imaginaries. While it is not possible to capture all media reports on a given problem domain such as sustainable cities world-wide, a range of reports from a variety of journals and industry reports should be consulted, ensuring that journals reflect a variation in political views, ownership, and target audience. A second data set useful for the same type of analysis is government press releases and public speeches. To focus searches, a set of keywords is used that emerge from my evaluation framework along with looking for mentions of the process explicitly and the prototype projects developed by participants. This data set will also enable the researcher to reflect on of societal effects so as to understand those that might be attributed to the process and those not. Computer assisted content analysis can also be used to assist in the environmental scan. If the right media databases exist, one could indeed "capture all media reports on a given problem domain such as sustainable cities worldwide." This would also help sort out the effects of the MUF versus those of exogenous events like economic fluctuations, local elections, global trade and climate negotiations, etc. In addition to data from interviews and bellwethers, tools such as issuecrawler or googlescraper which "enable the identification of diverse actors and collectives that have become publicized and make up an issue-space." (Chilvers & Kearnes, 2016, p. 273). If there is time and budget available, a social media analysis of the participants and the bellwethers can be conducted. This would provide a useful triangulation of what they are saying in public compared to what was said in interviews and provide an ongoing insight into the evolving public conversation on energy system transition.

Acknowledgments

I would like to thank the Mistra Urban Futures staff and LIP Directors for their generous feedback on earlier versions of this report. I would also like to thank Dr. John Robinson for his input and thoughts on implementing the framework.

References

Abelson, J., Forest, P.-G., Eyles, J., Smith, P., Martin, E., & Gauvin, F.-P. (2003). Deliberations about deliberative methods: issues in the design and evaluation of public participation processes. *Social Science & Medicine*, *57*(2), 239–251. http://doi.org/10.1016/S0277-9536(02)00343-X

Baldwin, S. (2000). Interactive social science in practice: New approaches to the production of knowledge and their implications. *Science and Public Policy*, *27*(3), 183–194. http://doi.org/10.3152/147154300781782048

Barrett, G., Wyman, M., & Coehlo, V. S. P. (2012). Assessing the policy impacts of deliberative civic engagement. In T. Nabatchi, J. Gastil, M. Leighninger, & G. M. Weiksner (Eds.), *Online deliberation design: Choices, criteria, and evidence*. Oxford University Press. http://doi.org/10.1093/acprof:oso/9780199899265.001.0001

Black, L., Burkhalter, S., Gastil, J., & Stromer-Galley, J. (2008). Methods for Analyzing and Measuring Group Deliberation. In L. Holbert (Ed.), *Sourcebook of Political Communication Research Methods, Measures and Analytical Techniques*. New York.

Blackstock, K. L., Kelly, G. J., & Horsey, B. L. (2007). Developing and applying a framework to evaluate participatory research for sustainability. *Ecological Economics*, *60*(4), 726–742. http://doi.org/10.1016/j.ecolecon.2006.05.014

Bremer, S., & Meisch, S. (2017). Co-production in climate change research: reviewing different perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, *38*(Suppl. 3), e482–22. http://doi.org/10.1002/wcc.482

Buick, F., Blackman, D., O'Flynn, J., O'Donnell, M., & West, D. (2016). Effective Practitioner–Scholar Relationships: Lessons from a Coproduction Partnership. *Public Administration Review*, *76*(1), 35–47. http://doi.org/10.1111/puar.12481

Caddy, J. (2005). Evaluating public participation in policy-making. OECD Publishing.

Carpini, M., Cook, F. L., & Jacobs, L. R. (2004). Public deliberation, discursive participation, and citizen engagement: A review of the empirical literature. *Annual Review of Political Science*, *7*(1), 315–344. http://doi.org/10.1146/annurev.polisci.7.121003.091630

Cornish, L. (2013). *Can 4D visioning foster community responses on climate change?*

Currie, M., King, G., Rosenbaum, P., Law, M., Kertoy, M., & Specht, J. (2005). A model of impacts of research partnerships in health and social services. *Evaluation and Program Planning*, *28*(4), 400–412. http://doi.org/10.1016/j.evalprogplan.2005.07.004

Fazey, I., Bunse, L., Msika, J., Pinke, M., Preedy, K., Evely, A. C., et al. (2014). Evaluating knowledge exchange in interdisciplinary and multi-stakeholder research. *Global Environmental Change*, *25*, 204–220. http://doi.org/10.1016/j.gloenvcha.2013.12.012

Fiester, L. (2011). *Measuring Change While Changing Measures: Learning In, and From, the Evaluation of Making Connections*. Baltimore, MD: The Annie E. Casey Foundation.

Filipe, A., Renedo, A., & Marston, C. (2017). The co-production of what? Knowledge, values, and social relations in health care. *PLOS Biology*, *15*(5), e2001403–6. http://doi.org/10.1371/journal.pbio.2001403

Gaventa, J., & Barrett, G. (2010). So What Difference Does it Make? Mapping the Outcomes of Citizen Engagement. *IDS Working Papers*, *2010*(347), 01–72. http://doi.org/10.1111/j.2040-0209.2010.00347_2.x

Hansson, S., & Polk, M. (2017). *Evaluation of knowledge co-production for sustainable urban development*. Mistra Urban Futures.

Kasemir, B., Jager, J., Jaegar, C. C., & Gardner, M. T. (Eds.). (2003). Public Participation in

Sustainability Science: A Handbook. Cambridge: Cambridge University Press.
Klein, J. T. (2008). Evaluation of Interdisciplinary and Transdisciplinary Research. *American Journal of Preventive Medicine*, *35*(2), S116–S123. http://doi.org/10.1016/j.amepre.2008.05.010
Koch, J. (2014). *Outcome Harvesting: An Overview of What it is and How it Can be Used*. Calgary: Community Development Learning Initiative.
Luederitz, C., Schapke, N., Wiek, A., Lang, D. J., Bergmann, M., Bos, J. J., et al. (2016). Learning through evaluation – A tentative evaluative scheme for sustainability transition experiments. *Journal of Cleaner Production*, 1–42. http://doi.org/10.1016/j.jclepro.2016.09.005
Mayne, J. (2001). Addressing Attribution Through Contribution Analysis: Using Performance Measures Sensibly. *The Canadian Journal of Program Evaluation*, *16*(1), 1–24.
Mistra Urban Futures. (2015). Mistra Urban Futures Strategic Plan 2016–2019, 1–62.
Mutz, D. C. (2008). Is Deliberative Democracy a Falsifiable Theory? *Annual Review of Political Science*, *11*(1), 521–538. http://doi.org/10.1146/annurev.polisci.11.081306.070308
Pincock, H. (2012). Does deliberation make better citizens. In T. Nabatchi, J. Gastil, M. Leighninger, & G. M. Weiksner (Eds.), *Online deliberation design: Choices, criteria, and evidence*. Oxford University Press. http://doi.org/10.1093/acprof:oso/9780199899265.001.0001
Quinlan, K. M., Kane, M., & Trochim, W. M. K. (2008). Evaluation of large research initiatives: Outcomes, challenges, and methodological considerations. *New Directions for Evaluation*, *2008*(118), 61–72. http://doi.org/10.1002/ev.261
Robinson, J., & Tansey, J. (2006). Co-production, emergent properties and strong interactive social research: the Georgia Basin Futures Project. *Science and Public Policy*, *33*(2), 151–160. http://doi.org/10.3152/147154306781779064
Robinson, J., Burch, S., Talwar, S., O'Shea, M., & Walsh, M. (2011). Envisioning sustainability: Recent progress in the use of participatory backcasting approaches for sustainability research. *Technological Forecasting & Social Change*, *78*(5), 756–768. http://doi.org/10.1016/j.techfore.2010.12.006
Rowe, G., & Frewer, L. J. (2005). A typology of public engagement mechanisms. *Science, Technology & Human Values*. http://doi.org/10.1177/0162243904271724
Ryfe, D. M. (2005). Does Deliberative Democracy Work? *Annual Review of Political Science*, *8*(1), 49–71. http://doi.org/10.1146/annurev.polisci.8.032904.154633
Simon, D., Arfvidsson, H., Anand, G., Bazaz, A., Fenna, G., Foster, K., et al. (2016). Developing and testing the Urban Sustainable Development Goal's targets and indicators – a five-city study. *Environment and Urbanization*, *28*(1), 49–63. http://doi.org/10.1177/0956247815619865
Talwar, S., Wiek, A., & Robinson, J. (2011). User engagement in sustainability research. *Science and Public Policy*, *38*(5), 379–390. http://doi.org/10.3152/030234211X12960315267615
The World Bank. (2014). *Cases in Outcome Harvesting*. Washington, DC.
Wall, T. U., Meadow, A. M., & Horganic, A. (2017). Developing Evaluation Indicators to Improve the Process of Coproducing Usable Climate Science. *Weather, Climate, and Society*, *9*(1), 95–107. http://doi.org/10.1175/WCAS-D-16-0008.1
Walter, A. I., Helgenberger, S., Wiek, A., & Scholz, R. W. (2007). Measuring societal effects of transdisciplinary research projects: Design and application of an evaluation method. *Evaluation and Program Planning*, *30*(4), 325–338.

http://doi.org/10.1016/j.evalprogplan.2007.08.002
White, H., & Phillips, D. (2012). *Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework* (No. Working Paper 15). International Initiative for Impact Evaluation.
Wiek, A., Talwar, S., O'Shea, M., & Robinson, J. (2014). Toward a methodological scheme for capturing societal effects of participatory sustainability research. *Research Evaluation*, *23*(2), 117–132. http://doi.org/10.1093/reseval/rvt031
Williams, S. (2017, June). *Evaluating Transition Experiments in Times of Rapid Change*. Paper presented at International Sustainability Transitions 2017. Chalmers University, Gothenburg, Sweden.

Appendix A: Integrated Evaluation Framework

Order	Category	Effect	Proposed Indicators	Research Methods
1st	Usable Products	Technologies and social innovations Products/services Action plans Publications (academic & non-academic) Distribution of knowledge	Innovative technologies and social innovation adoption/scale Intellectual property – innovation and new ideas Action plans Media – newspaper, magazine, online articles, reports Written reports, research and practice overviews, (spreading knowledge) meetings with decision makers, workshops, seminars; funding applications; peer reviewed articles; technical reports/gray literature # times results of the programme were the subject of public or private discussions of the participant (apart from the transdisciplinary process), # references to the programme were made by others in his presence, # references the participant himself had made to his fellow citizens.	**Data collection:** Meeting observation Prototype tracking Document review Semi-structured interviews with participants **Analysis:** Actor Network Tracing Outcome harvesting
	Enhanced Individual Capacity	Acquired knowledge (individual or collective) Understanding (vocabulary, perspective, etc.) Organizational learning Know-how of technologies Anticipatory competence Process knowledge Personal change Understanding Systems Knowledge Decision Making Capacity	Co-produced knowledge System, goal or transformation knowledge Co-produced scenarios Idea exchange Transformation of knowledge [also in policy?] Changes in understanding - increased knowledge, change in attitudes, changes in thinking Changes in ways of thinking, learning and working New understandings of subject area (e.g. practices of urban planning) [potential new category of knowledge exchange/distribution – e.g. provision of information] Changed perceptions and expectations (i.e. Understanding of perspectives and preferences of different stakeholder groups); broadening of perspectives and knowledge Increased understanding of desires and concerns of collaborators Enhanced communication skills Awareness of oppositional arguments Attitude and behaviour change Changes in roles Greater sense of empowerment and agency	

Order	Category	Effect	Proposed Indicators	Research Methods
			New skills, increased confidence Insight into own role in system, understanding of problem context	
	Network Effects	Networks created or expanded Community trust created or expanded Trust Accountability Community identity	Expanded participant networks Boundary-crossing collaborations and partnerships New contacts Strengthened (existing) networks # new people met, , participated in further working meetings, visited public meetings on related topics, professional meetings See National, international and local partnerships (see strategic plan pp. 33-38) and global work pp. 38-40 Willingness share with broader networks New ways of working across disciplines and sectors; better understandings of roles of partner orgs; better understanding of how to make use of expertise and networks; collaborative funding applications; network building, community identification Willingness to work on a specific problem; readiness to share knowledge with fellow citizens, to intensify co-operation, and to leave important tasks in a joint project to others Social capital / feelings of identity	
Sources	Wiek et al. 2014; Luederitz et al. 2016; Riddell 2015; Gaventa & Barrett 2010; Mutz 2008; Ryfe 2005; The Natural Step 2017b; Nelson & Nelson 2002; Hanssen & Polk 2017; Fazey et al. 2014; Reepalu et al. 2015; Walter et al. 2007			

Order	Category	Effect	Proposed Indicators	Research Methods
2nd	Policy Effects	Economic benefits Policies Decisions made Solutions implemented Instrumental use Conceptual use	Resource management decisions Policies/Laws passed Change policy discourse; Effects on discourse e.g. new perspectives and questions on the agenda New evidence introduced into policy/strategy; synthesis and accessibility of information to policymakers; Distribution of knowledge; informing management decisions, policy actions or adaptation decisions Direct impact on policy and practice decisions e.g.– conceptual (better informed), justification (justifying a predetermined decision), instrumental (inform a new decision) Indirect impact on the knowledge, understanding and attitudes of policy-makers and practitioners Infrastructure changes Changes in policy or practice leading to next dimension of evaluating *impacts* of changes in policy or practice – e.g. ecological health, social and economic welfare, social equity, business performance, etc.	**Data collection:** Bellwether interviews Semi-structured interviews with participants Organizational interviews **Analysis:** Outcome harvesting Process tracing Most significant change Contribution analysis
	Organizational Changes and Actions	Changed context for new and ongoing work New organizations New business models New standards New partnerships	Shift in org responsibilities/roles Changes in job description Shift in rules of engagement New institutional frameworks Shifts in investment strategy Changes in operating models Programme results used in teaching materials Change in decision making process	
Sources	Wiek et al. 2014; Barrett et al. 2012; Meagher 2008; The Natural Step 2017b			

3rd	Alternative Visions and Imaginaries	Shifts in public narrative Collective purpose and vision Greater social cohesion across groups	New public discourse Activism Artistic engagement Resiliency of new narratives	**Data collection:** Context monitoring Semi-structured interviews with participants Semi-structured interviews with design team and partners Bellwether interviews **Analysis:** Process tracing Most significant change Contribution analysis
	Transformed Social Practices	Landscape shift Norm change and/or adoption Inclusion of new actors and issues in public spaces and discourse Changes in practices of participation New space for innovation and experimentation	Behavioural change Increased capacities for collective action New forms of participation Deepening of networks and solidarities	
Sources	Chilvers & Kearnes 2016; Taylor 2004; Bichieri 2017; Gaventa & Barrett 2010; The Natural Step 2017b; Hanssen and Polk 2017; Wall et al. 2017; Walter et al. 2007; Fazey et al. 2014; Oh 1996			