A FRAMEWORK TO EVALUATE LOW-CARBON ENERGY TRANSITION LEARNING PROJECTS

Internal report presented to Natural Resources Canada

by scholars and collaborators of

Sustainable Canada Dialogues

September 2018

Authors: Potvin, C.[a], Sharma, D.[a], Williams, S.[b], Ashley, B.[c], Hall, M.[d], Hoffmann, M.[e], McVey, I.[f], Meadowcroft, J.[g], Morency, C.[h] Robinson, J.[e], Rosenbloom, D.[g], Sheppard, S.[i], Vasseur, L.[j] and Wielinski, G.[h]

[a] Department of Biology, McGill University, 1205 Dr Penfield Ave., Montreal, QC, H3A 1B1
[b] Faculty of Science, University of British Columbia, 429 – 2202 West Mall, Vancouver, BC, V6T 1Z4
[c] City of Summerside, 275 Fitzroy St, Summerside, PE, C1N 1H9
[d] Faculty of Sustainable Design Engineering, University of Prince Edward Island, 550 University Avenue, Charlottetown, PEI, C1A 4P3
[e] Munk School of Global Affairs, University of Toronto, 1 Devonshire Place, Toronto, ON, M5S 3K7
[f] Ontario Climate Consortium Secretariat, Toronto And Region Conservation Authority, 101 Exchange Avenue Vaughan, ON, L4K 5R6
[g] School of Public Policy and Administration, Carleton University, 5224 Richcraft Hall, Ottawa, ON, K1S 5B6
[h] Department of Civil, Geological and Mining Engineering, Polytechnique Montréal, 2900, boul. Édouard-Montpetit, Montréal, QC, H3T 1J4
[i] Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC, V6T 1Z4
[j] Department of Biological Sciences, Brock University, 1812 Sir Isaac Brock Way, St. Catharines, ON, L2S 3A1

Table of Contents

1. ACKNOWLEDGEMENTS

We would like to thank the following collaborators for their participation in a workshop held on June 1, 2018, and for their feedback on the document as the proposed evaluation framework took shape: Frédéric Beauregard Tellier, Environment and Climate Change Canada; Nicholas Chesterley, Employment and Social Development Canada; Ian Clark, NRCan; Meaghan Davis, City of Toronto; Dror Etzion, McGill U., SCD; David Hughes, The Natural Step Canada; Michael Kalin, NRCan; John Kenney, NRCan; Haris Khan, Privy Council Office; Mary Kay Lamarche, NRCan; Niamh Leonard, McConnell Foundation; Drew Leyburne, NRCan; Mishka Lysack, U. Calgary, SCD; Deepti Mathew Iype, U. British Columbia, SCD; Damon Matthews, Concordia U., SCD; Patrick McCurdy, U. Ottawa, SCD; Janetta McKenzie, U. Waterloo, SCD; Dan Monafu, Treasury Board Secretariat; Normand Mousseau, U. de Montréal, SCD; Eric Poirier, Infrastructure Canada; Shahrzad Rahbar, Industrial Gas Users Association; Caroline Sanchez Valero, Réseau Environnement; Mark Stoddart, Memorial U., SCD. We are most grateful to the practitioners who took the time to discuss their visions of an evaluation framework, from the City of Summerside: Gordon MacFarlane; Greg Gaudet; Aaron MacDonald; Gerald Giroux; Mike Thususka; Rob Philpott; Bobby Dunn; and the Toronto and Region Conservation Authority: Sonya Meek; Adriana Gomez; Shannon Logan; Jose Manuel Torcal.

2. CONTEXT

Federal Budget 2018 included a focus on making progress towards fulfilling Canada's commitments to reduce greenhouse gas (GHG) emissions and transition to a low-carbon economy. This transition will require the implementation of a mix of different low-carbon energy sources and low-carbon energy usage to meet national energy demands. What the 'made-in-Canada' low-carbon energy mix will look like is unknown and the potential role for each energy source and usage needs to be explored. Informed policy is essential to drive change from business-as-usual scenarios in a way that addresses low-carbon energy goals, but also maximizes economic benefits, maintains Canada's competitiveness, ensures reliability and considers environmental and social impacts.

The purpose of this report is to develop a framework to evaluate socio-technical energy **Transition Learning Projects**. It builds on recommendations from *Re-Energizing Canada: Pathways to a Low-Carbon Future*, a report commissioned to Sustainable Canada Dialogues (SCD) by Natural Resources Canada (NRCan) in the context of Generation Energy. One of the SCD scholars' recommendations for immediate action is to support low-carbon energy transition doing-by-learning projects in innovative social practices and technologies covering the spectrum of diversity found in Canada. The present report synthesizes the literature on low-carbon energy transitions and their evaluation. It then proposes an evaluation framework developed and enriched by discussions and exchanges that took place at a one-day working session held in Ottawa on June 1, 2018, when SCD invited a group of experts with relevant experience from the private, NGO and public sectors to begin co-developing a holistic approach for evaluating socio-technical energy **Transition Learning Projects**. The proposed evaluation framework is illustrated by three case studies and benefited from several rounds of feedback from practitioners, federal civil servants and scholars.

3. LOW-CARBON ENERGY TRANSITION PROJECTS

3.1. Defining Transition Learning Projects

The low-carbon energy transition (hereafter the low-carbon transition) entails not only the diffusion of new and existing technologies (e.g., the adoption of electric vehicles and renewable energy sources) but also changes to policy and regulatory frameworks (e.g., regulatory standards and carbon pricing regimes) to facilitate innovation and speed up the adoption of novel social practices and norms (e.g., new perceptions of, and behaviours, around mobility and electricity use).[1,2,3] The transition is taking place in the context of rapid and potentially profound changes in the social and technical

systems in which these practices and norms occur, including changes in information and communications technology, settlement patterns, livelihoods and occupations, mobility, cultural beliefs and practices, and production and consumption systems and activities, all of which offer both opportunities and challenges for intentional low-carbon transition strategies. In the context of necessary transformational changes, widespread and diverse projects are essential, as we do not know which approaches will work best in particular circumstances.[4] The characterization of the costs and benefits of different approaches, their distribution over time and space, and their relative desirability and social acceptability cannot be fully known in advance. Learning, therefore, serves as a core component of processes of system innovation (see Box 1 for an international example), particularly the iterative relationship between 'doing' and 'learning' as transition projects begin to carry forward transformative change on the ground.[5,6]

Box 1. Finland's 'A Place to Experiment' digital platform as a model for promoting Transition Learning Projects

Finland has adopted a strategy to engender a culture of experimentation among government and societal actors.[a] This strategy aims to generate knowledge about the design of new legislative, institutional and operational models. To do so, the Government has created a digital platform, 'A Place to Experiment', to build capacity, coordinate feedback and learning, create networks between innovators and government, provide information on funding sources and promote innovations.[b] As part of this, Finland is moving to encourage further pilot projects and innovation around major societal challenges such as energy efficiency and climate change. Two experimentation hubs that have emerged within this context are the Carbon Neutral Municipalities project—a government facilitated network of municipalities, businesses and residents—and the smart district of Kalasatama in Helsinki—a brownfield development aimed at reimagining urban functions. This has led to the development of a portfolio of projects across sectors and at multiple scales, from larger projects exploring new forms of housing and energy provision to smaller citizen-driven initiatives tackling consumption practices.[c]

Transition Learning Projects therefore are: (a) *explicitly linked to a larger vision* of the long-term transformation of our energy systems, and (b) *consciously understood as pilots* to test the viability of innovations and explore possibilities for change.[7] They can be understood as deliberate interventions that experiment with a novel configuration of social and technical elements that could lead to substantial low-carbon change in energy systems.[8] Rather than simply demonstrating technologies or technological applications, the **Transition Learning Projects** would act as small-scale trials of new models for meeting societal functions, shedding light on novel ways to produce and distribute energy, manufacture products, transport people and goods, or develop built environments.

3.2. Benefits of Transition Learning Projects

There are four critical advantages that **Transition Learning Projects** can bring to efforts to realize a low-carbon energy transition. First, as mentioned above, they can enable *learning* about change, what does and does not work, thus helping to improve and scale up possible responses. Second, they can promote *capacity-building,* especially by joining up networks of innovators (e.g., policymakers, community groups and entrepreneurs). Third, they can assist in *de-risking* the innovation process, for example through partnerships, by pooling societal resources and expertise or by testing new policies and regulations (see Box 2). International engagement around transition experiments indeed indicates that there is as much to learn from 'failures' as there is from successes.[9,10,11,12,13,14,15] Fourth, they have enormous *educational potential* as a citizen outreach and engagement tool, thus fostering greater social acceptability for the low-carbon energy transition.

Box 2. UK Regulatory Sandbox as an innovation de-risking tool

In 2017, the UK Office of Gas and Electricity Markets initiated a 'regulatory sandbox' to foster real innovation in the UK energy sector, acknowledging that transformative change is a necessary component of the transition to low-carbon energy that can be hampered by current regulations.[d] To enable entrepreneurs to test innovative technologies and business models that seek to benefit consumers, barriers to innovation are temporarily removed or modified and safeguards applied to ensure that risks are not taken up by consumers. The Office's Innovation Link guides innovators through energy regulations and exemptions or modifications thereof.[d]

The literature suggests that learning projects are particularly useful when they combine both top-down and bottom-up approaches, are deployed at different scales and contexts, involve stakeholders from different societal sectors and link the attainment of climate- and energy-related objectives to the satisfaction of specific societal needs (including health and welfare benefits, improved service, functionality and convenience for consumers, enhanced community control or resource savings and expanded commercial opportunities).

3.3. Learning from Transition Learning Projects

What therefore distinguishes the proposed **Transition Learning Projects** from current innovations is an explicit orientation to evaluate project experiences, extract and share lessons learned and inform subsequent decision-making. The literature suggests that evaluating projects aimed at encouraging a low-carbon transition requires a comprehensive and iterative approach.[16] In this way, a framework for evaluating **Transition Learning Projects** should assess: impacts on multiple levels using quantitative and qualitative tools; crucially, how the project is contributing more broadly to a low-carbon transition;[17] and how and where the lessons learned could be shared and possibly replicated or scaled up. The iterative role of evaluation entails that it also serves as a learning tool for innovators who are conducting projects, for example allowing participating firms to identify capacity gaps around novel business models (e.g., energy as a service) or new user preferences, and to bridge these gaps to take advantage of emerging market opportunities. This in turn fosters critical thinking, self-awareness of capacity limitations, systemic weaknesses and embedded bias.

These aspects of learning from transition projects are well-illustrated by the Alberta Energy Futures Lab (EFL).[18] This multi-stakeholder engagement project has a goal of supporting a transition to a sustainable energy system in Alberta where the political, social and economic context has changed dramatically over the past five years. In the EFL, evaluation has been used to help innovators understand the effects of the project on participants and the energy system and to help refine strategy choices for the EFL design team. For example, evaluators mapped the portfolio of EFL initiatives against a visual representation of the energy system in Alberta, a strategic foresight framework and sources of GHG emissions in the province.[19] Insights from this analysis led to program changes such as the inclusion of representatives from the agriculture and transportation sectors along with expanding the EFL portfolio to include a wider range of initiatives than just technical solutions, including cultural engagement projects.

4. CO-DEVELOPING AN EVALUATION FRAMEWORK

4.1. Method

To co-develop a framework for evaluating socio-technical energy transition experiments, SCD collaborated in a number of ways with experts with relevant experience from the private, NGO and public sectors (Figure 1).

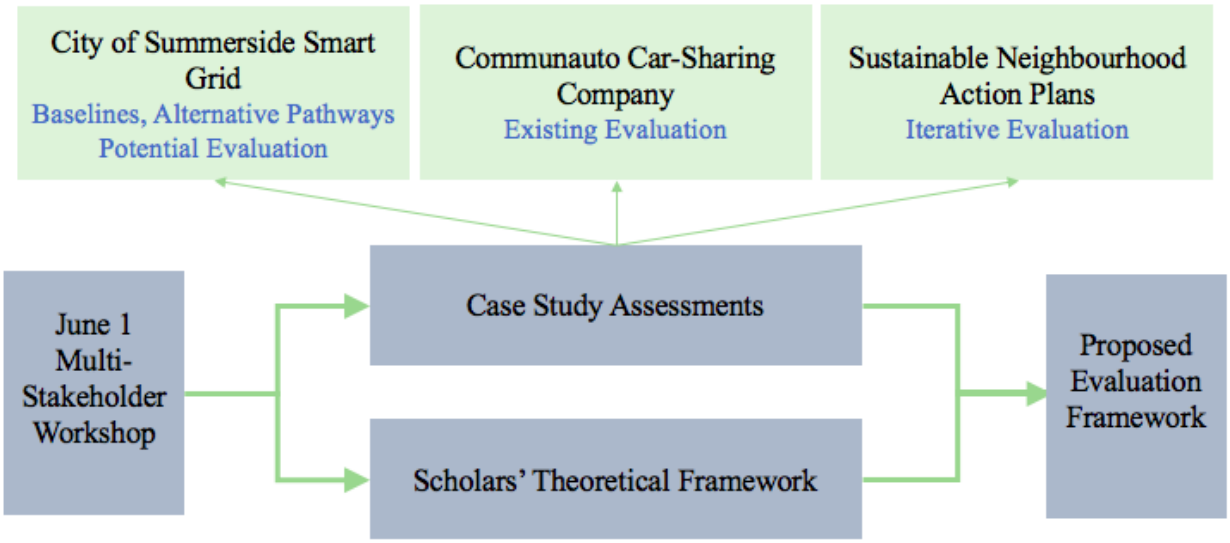

Figure 1. Participatory process over Summer 2018 to develop a framework to evaluate Transition Learning Projects.

The work began on June 1, 2018, when 27 stakeholders from academia, municipal and federal government, industry and NGOs participated in a workshop held in Ottawa to identify desired elements of a framework to evaluate **Transition Learning Projects**. There was broad, cross-sector agreement on (1) the need for an evaluation framework that focuses on transformative social and environmental change that enables a sustainable transition towards low-carbon energy, and (2) the importance of scaling up learning and motivating behavioural change (see Appendix 1 for full summary).

Participants felt that the framework should:

- Apply at the onset, as early as possible during project design, and over the entire course of the learning project. The framework should thus allow reflexive adaptation/course-correction/feedback as the external context changes and as we learn from the project, evaluating the value created by the process as well as the outcomes.
- Be flexible and general enough to apply to various kinds of learning projects with various definitions of success, but also specific enough to be relevant to and compare particular projects. For example, it was suggested that the evaluation framework needs to be broad enough such that, even if only qualitative evidence and case studies are used as methods, evaluation could still be approached with the proposed framework in mind.
- Identify actionable recommendations and results that could be replicated, scaled up and generalized to enable learning between and from learning projects, to raise public awareness, encourage engagement and behaviour change and build long-term capacity, and thereby accelerate the low-carbon transition. The framework should therefore aim to attribute effects to causes (i.e., attribution as compared to contribution to effects), teasing out what can be applied in other contexts versus what is context-dependent.
- Be a multi-stakeholder collaboration that is participatory/co-created with project proponents who might not have the capacity to undertake the evaluation themselves, and with eventual users of the framework. This approach helps build relationships, share ideas and enable dialogue between stakeholders, who begin the transition project with an understanding of shared values and visions.
- Be simple, practical and transparent, in order to be applied effectively.

While all participants agreed that the evaluation framework had to be rigorous, some stated that the best way to do so was to compare the intervention with a baseline and/or control group. Participants discussed how to translate the desire, expressed in particular by federal government participants, for the "gold standard" of experimentation—as in, randomized control trials—to the complex real-world systems of transition projects where such controlled settings are often not applicable (see, e.g., [20,21]). Traditional experimental thinking understands that comparison is to a so-called baseline or some other control group that functions as a counterfactual. In the language of experimentation, the counterfactual is used to tell us what would have happened otherwise.

The scholars then produced an initial conceptual framework, using a literature review to shed light on the topics discussed in the workshop. For example, the literature suggests that, given the complexity of the multiple social-ecological-political-economic systems at play, comparing to a baseline or counterfactual is not always possible[22,23,24] nor desirable, when there is no meaningful baseline (see [25] for a thorough discussion on how to understand experimentation as interventions versus controlled trials).

Simultaneously, another set of scholars engaged with case study practitioners to understand what constitutes an evaluation framework useful to project proponents. Specifically, we interviewed practitioners from the City of Summerside, PEI—which has a smart grid project with no formal evaluation—and the Toronto and Region Conservation Authority—where an iterative evaluation has been initiated for its Sustainable Neighbourhood Action Plan. We then laid out the Québec car-sharing company Communauto's evaluation to elucidate how an existing evaluation can fit into the proposed framework. The scholars' and practitioners' thinking on evaluation was used to co-create the proposed framework.

4.2. Literature on Evaluation of Socio-Technical Energy Transitions[26]

The sustainability transition literature is relevant to the development of an evaluation framework for low-carbon energy **Transition Learning Projects**. This literature emphasizes the need for managing transition projects for deepening (i.e., learning), broadening (i.e., replicating) and scaling up (i.e., promoting transformative innovation) as mentioned in Section 3.3 above.[27] Across the transition literature, a desirable evaluation is regarded as flexible, continuous or long-term (occurring before, during and after the project), generic (i.e., comparable and usable across different kinds of projects) and comprehensive.

Evaluating socio-technical energy **Transition Learning Projects** poses a number of challenges beyond those of traditional evaluation. As numerous authors have pointed out, attempting to foster a sustainability transition entails working within complex systems.[28,29,30,31,32] Not only is the system constantly changing—which poses challenges of attribution—system transitions, especially energy system transitions, tend to take place over long time periods.[33,34,35] Capturing the long-term contribution of an intervention that runs for months or even years to system transition is very difficult; even more so given that, due to finite resources, evaluations may only focus on a subset of the project lifecycle. A third challenge is the boundary-spanning nature of transitions across multiple domains (e.g., social, political, cultural and technical).[36] Developing tools and methods to capture change across such wide-ranging domains is difficult both conceptually and practically. Finally, there is a tension between the need for evaluation that generates learning for those designing and managing sustainability **Transition Learning Projects** and accountability for 'results' driven by funders and governments. A finding of "'we had a good process and learnt things' does little to address the need for accountability in government interventions".[37,38] This tension is especially acute for transition projects that have a specific goal of reducing GHG emissions in a given sector, geographic area or target population, for example.[39]

Literature on evaluation of participatory and public engagement processes also provides insights for evaluating low-carbon transition projects. Such evaluations are important because they 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 and fair and present unbiased information.[40,41,42] While measures of process effectiveness are well-established, outcome measures in the context of low-carbon transitions are much less developed. Outcome evaluation looks at impact of the process and the products it produces (e.g., reports) on target audiences (e.g., in terms of increased knowledge or level of civic participation) as well as the use of those products by decision-makers.[43,44,45,46,47,48] A commonality of this literature is that it comes from a governance perspective, asking how engaging with citizens will link to policy- and decision-making.[49]

Equally relevant, an evaluation framework for capturing societal effects of engagement processes has emerged[50,51,52] from the literature on the impacts of participatory research.[53,54,55,56,57,58,59] 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, websites, 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 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.[60] These effects include changes in energy use and emissions, whether immediate or emerging over time, crucial to meeting official quantitative emission targets and to the low-carbon transition.[61,62] Beyond these first- and second-order effects, learning projects can also lead to new, alternative visions of the future, for example through "forms of activism and artistic engagement", and broader social transformations, for example via "pro-environmental behaviour change initiatives".[63] This set of effects can be thought of as a third order—how societies conceive of themselves and behave—which can include effects such as norm and behaviour change. One can then attempt to measure impacts of transition projects, asking whether they strengthened socio-ecological integrity, enhanced livelihood sufficiency and opportunity and other measures of sustainability.[64] However, addressing how a given process is contributing to a sustainability *transition* is distinct from evaluating specific indicators of sustainability. To do so, evaluation can also explicitly consider "core societal effects" of sustainability transition management processes: social learning, empowerment and social capital.[65]

Another avenue to develop an evaluation framework is theory-driven evaluation, traced to the early 1930s.[66] Such evaluation framework has been referred to more or less interchangeably as "program-theory evaluation, theory-based evaluation, theory-guided evaluation, theory-of-action, theory-of-change, program logic, logical frameworks, outcomes hierarchies, realist or realistic evaluation". Because transition thinking entails triggering socio-technological and behavioural changes, theory-driven evaluation provides a structure to guide the development of the proposed evaluation framework. In such evaluations, what success looks like is identified upfront—in government, this is typically through a model and its associated indicators. Indeed, the theory of change is a tool that allows a description of how and why activities carried out by a project will elicit the desired changes.[67] As a result, the evaluation emerges from the process of change that is identified at the onset and that motivates the proposed activities.

Early linear theory-of-change models were built around inputs, activity, outputs and outcomes, the latter being perceived as short-, middle- and long-term. Outputs and outcomes differ mainly in that the former result directly from an intervention and are controlled by project proponents, while the latter are often outside of proponents' control.[68] A simple example is a program designed to build awareness or knowledge. The *outputs* would be the number of training sessions, training materials

prepared, and so on; but an *outcome* would be increased awareness or increased capacity—i.e., knowledge, skills and abilities. The theory-of-change model that links these different components lays out assumptions of how the project is intended to bring about the desired changes.[69] A strength of theory-of-change evaluation therefore is that it examines whether the inputs and activities serve to make progress not only towards outputs, like in traditional evaluation, but also towards the outcomes identified vis-à-vis the intended objectives of the project.[70]

More recent theory-driven evaluation models have moved away from the linear relationship to integrate systems thinking, recognizing that evaluation should consider all contexts (societal, economic, political, and so on), be comprehensive and take in to account external and internal factors that could influence outputs and outcomes.[71]

5. AN EVALUATION FRAMEWORK PROPOSAL

Based on the literature and June 1 workshop results, the scholars developed an initial conceptual evaluation framework. The framework was then refined following feedback from case study practitioners and June 1 stakeholders (Figure 2). The framework is theory-driven and, as such, the proposed method is familiar to Natural Resources Canada, which used theory-of-change models for, *inter alia*, the evaluation of their Climate Change Adaptation[72] and Energy Efficiency Sub-Programs.[73] Expertise thus exists within Natural Resources Canada to fully implement our proposal.

Theory-driven evaluation considers assumptions at play, where the arrows in a model represent expectations about what conditions must be in place for the causal relationship to exist. Here, we propose to organise our theory-of-change model in a way that differentiates projects' 'results' into three categories: output, outcome and impact.

Short-term outputs are the tangible, short-term products of a project that are under the proponents' direct control and where the immediateness of the effect can be easily monitored.

Medium-term outcomes are medium-term effects of a project that relate back to the project goals but might be under less direct control of the proponents. It has been said that almost all projects aiming at social changes will do so through intermediate outcomes, such as improving services to users or modifying users' attitudes, behaviours or thinking, and so on. The causal relationships leading to medium-term outcomes therefore need to be clearly articulated in the theory-of-change model.

Impacts are the longer-term consequences of a project and can include tangible improvements in reducing GHG emissions resulting from a large-scale shift to low-carbon energy and longer-term changes to organizational culture, consumer behaviour or public attitudes and values. "*Impact*" attempts to capture how projects are fostering a transition to new development pathways or, in other words, moving towards a low-carbon economy. For example, transition theory suggests that changes in public narrative and values are markers of regime transition. Whether those narratives, values and subsequent behaviour choices are becoming more pro-environmental is an indicator of a transition in a more sustainable direction. Similarly, changed or newly established dominant social practices—e.g., practices of mobility, food choice and preparation, and housing—would indicate that sustainable practices are becoming the norm.

We propose this distinction between output, outcomes and impact because, while *long-term impacts* can be the most challenging to capture due to the lengthy time lag between an initial project and its impact, it is precisely these impacts that are the most important with respect to the low-carbon transition. If possible, measures or proxy measures should also be identified; at the least, why/how the short-term and medium-term outcomes will contribute to this impact should be clearly identified or articulated. Without a continual view to whether projects are contributing to long-term impacts, it is easy to fall back to projects that generate measurable short-term outcomes but do not contribute to

a societal transition to a low-carbon future. We believe therefore that purposefully acknowledging impacts in the evaluation framework will stimulate project proponents to consider elements of co-learning and scaling up for transformative change.

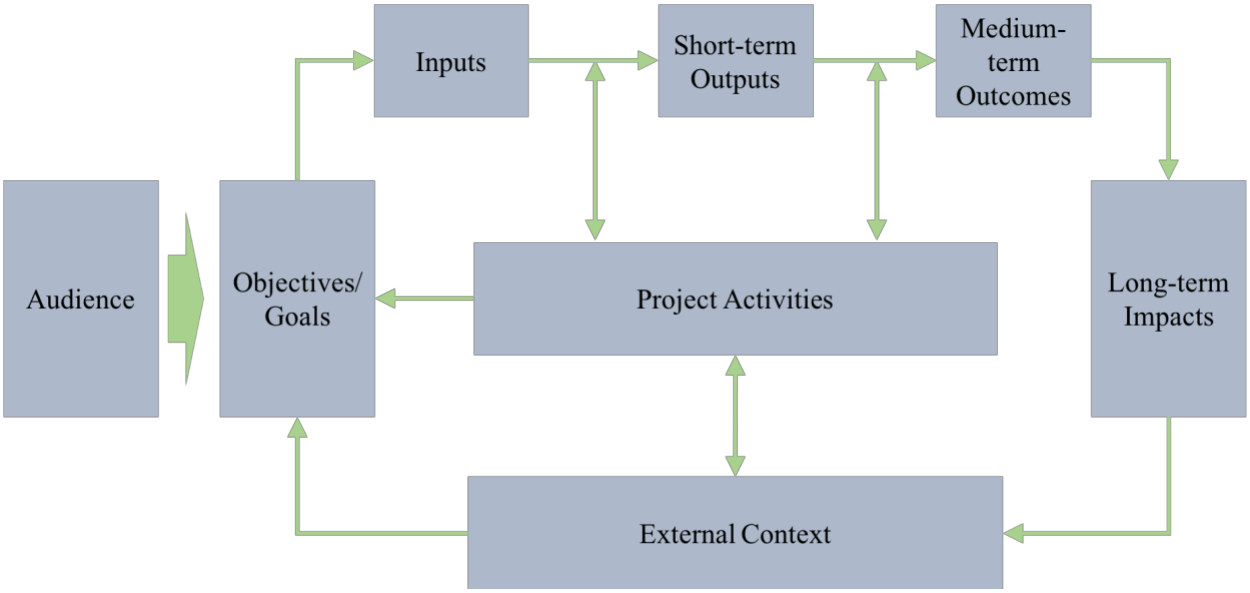

Figure 2. Generic theory-of-change model allowing evaluation of low-carbon Transition Learning Projects.

Adapted from Arnold, E., T. Åström, C. Glass, and M. De Scalzi. 2018. How should we evaluate complex programmes for innovation and socio- technical transitions? Brighton, UK: Technopolis Group. P. 18.

The proposed theory-of change model explicitly considers three contextual influences on **Transition Learning Projects**: (1) the ***audience*** of the project itself, a crucial first question that emerged from practitioner feedback. Both the initial development of the theory-of-change model and the subsequent evaluation procedures may be devised narrowly to engage certain stakeholders or broadly to engage multiple stakeholders. As **Transition Learning Projects** aim at mobilizing social changes, the latter should be favoured when possible; (2) ***external context*** factors, such as political context, socio-economic factors and characteristics of the target population, which can all have great consequences on individual projects and so should be carefully considered. Furthermore, the importance of local/regional external context factors is often undervalued when attempting to replicate the success of a project in a different context; (3) ***project activities*** (implementation characteristics), including the tools and methodology adopted by proponents, which can include the type of stakeholder engagement and methodologies used in the project design (e.g., alternative treatments/controls, nature of the interventions) and analysis of what triggers the outcome (e.g., were the face-to-face surveys most effective?).

An important caveat to consider when implementing the evaluation framework is that **Transition Learning Projects** need to be understood as working within complex systems in which the audience, the process and methodology chosen by the proponents, and the external factors all affect the outputs, outcomes and impacts. Accordingly, different stakeholders could consider distinct outputs or outcomes. Project proponents should be explicit about which target group or population they are trying to influence change in. For a city official evaluating a home retrofit **Transition Learning Project** to assess its scalability, for example, a policy could be an output, reserving the term outcomes for much broader public consequences. On the other hand, an NGO assessing the same home retrofit **Transition Learning Project** might consider municipal policy an outcome in an effort to promote

home retrofitting province-wide (see Case Study 2). Clarity with respect to the causal relationship in the theory-of-change model should help alleviate such discrepancies, as multiple stakeholders agree on the likely transition pathway described by the model. Furthermore, when experimenting within complex systems, the process is often characterized by iteration, feedback loops between system components and systems that are continually changing. Thus, transdisciplinary methodologies as well as quantitative and qualitative data are important to fully capture the changes occurring.

Critical to the outcome of an evaluation are the measures of success. In the theory-driven evaluation framework, since the output, outcome and impact indicators are determined *a priori*, the pathway of causal relationships that they describe becomes the main indicator of success. The development of an appropriate theory-of-change model is thus the first critical step in evaluation. It has been suggested that building a narrative theory of change is an efficient way of verifying the assumptions laid out in the model[74] (see Section 6 below). In early-stage **Transition Learning Projects**, a developmental evaluation[75] approach may be helpful. As opposed to formative and summative evaluation, which assumes that project success criteria and methods are well-defined and readily evaluable, developmental evaluation can help project proponents clarify their tentative theory-of-change model and, through a process of learning-by-doing, better articulate project goals and intended outcomes. This is an important consideration for social innovations working in complex systems where a linear path to "success" is not clear and experimentation is needed to better understand both the problem and the solution.

We recommend that project-specific indicators be chosen after the theory-of-change model is reviewed in a participatory way by key stakeholders. The indicators should measure steps towards achievement of the project's objective—i.e., contribute to the project's success based on the audience, inputs, external context and project activities. The evaluation should determine whether the desired values of indicators were achieved, whether the causal links between each category in the framework did occur as expected and to what extent, and whether and where unintended side effects and breakdowns in the expected relationships occurred.

As discussed, the complex nature of transition processes means that direct linear relationships between outputs, outcomes and impacts may be difficult to trace; evaluation should pay attention to feedback loops between outcomes and the process of the project itself. We recommend that projects use the generic theory-of-change model that we propose (Figure 2) to build their evaluation framework, including indicators that capture the causal relationships, i.e., pathways of change. The framework is thus amenable to different kinds of projects, as it leaves the determination of specific indicators to be decided by the evaluators depending on the project's objectives and theory-of-change model.

6. CASE STUDIES

Over Summer 2018, the scholars worked with practitioners to incorporate their visions for evaluation and feedback on the evaluation framework described above. Below we present three case studies in which the framework has been applied to an ongoing initiative that could be considered a **Transition Learning Project**. The case studies represent a range of possible low-carbon energy initiatives related to low-carbon energy use, mobility and energy efficiency. It is important to note that, in keeping with this report's objective, the purpose of working with the case studies was to allow us to refine the framework proposed rather than to evaluate the case studies themselves.

6.1. The City of Summerside Smart Grid Project

In 2009, Summerside Electric, the municipally-owned electric utility of the City of Summerside, Prince Edward Island, installed a 12 MW wind farm and contracted an additional 9 MW of capacity from another wind farm. As this constitutes a significant amount of wind energy relative to

Summerside's electricity demand, which peaks at 28 MW, Summerside Electric began a smart grid initiative to help make demand better match wind availability. These initiatives resulted in the notable achievement of wind energy providing 50% of the city's annual electricity consumption.[76]

Summerside's smart grid uses fibre optic and wireless communication to exchange real-time data between the utility and participating customers. This facilitates smart heating and hot water appliances which respond to grid signals. Summerside Electric's Heat for Less Now (HfLN) program sees residents and businesses replace existing heaters (typically oil furnaces) with electric thermal storage devices that combine resistance heaters with heat-storing ceramic bricks. The program also includes electric hot water heaters that can vary in temperature. Together, these appliances provide a means of adjusting electricity consumption by storing thermal energy, a capacity which is controlled by the utility in real time. Participation in the program is incentivized through discounts on the electricity consumed by connected appliances, provincial rebates and options to lease.[77] Participants receive feedback on appliance electricity consumption, costs and GHG emissions from an online customer portal. By shifting demand to time periods when more energy is available, the HfLN program provides financial benefits for both the utility and consumers.

To understand what the City would look for in a framework to evaluate its smart grid program, discussions were held with eight Summerside city officials, including senior management, in July 2018 (see full meeting summary in Appendix 2). Currently, evaluation of Summerside's smart grid project is *ad hoc*, infrequent and usually driven by a complaint or error, which involves sharing data or feedback, meetings to discuss and, at times, an assignment to look into alternatives or triggers like failures or deficiencies. Successes are sometimes also evaluated in terms of uptake. City officials consider evaluation valuable as a proof of concept for market validation; as a decision-making and forecasting tool—to know whether to pursue or modify the project; to help the City assess the effectiveness or efficiency of procured equipment, software and other supplies; as a toolkit for communications/outreach and public relations to both influence consumer behaviour with documented support, including political messaging and future policy development, and promote the project as participants share successes with non-participants; and as an integral part of grant funding applications.

Using this information, the following narrative theory of change was articulated for Summerside:

Context. A municipally-owned electric utility has substantial wind capacity that, at times, exceeds electricity demand. Its goal is to better match demand to periods of wind availability to avoid exporting the excess wind energy to the bulk grid at unfavourable prices.

Causal mechanisms. (1) Technological pathway: Implementing smart grid technology will align residential and commercial consumption of electricity with wind availability when electricity is less costly. (2) Communication pathway: Residents and businesses will be encouraged to participate in the smart grid program by communicating the financial and environmental benefits and peer-to-peer sharing of successes.

Working with the city officials, we also identified a range of indicators and applied them to the proposed evaluation framework (Figure 3). In the model of Summerside's theory of change, the two causal mechanisms and their associated outputs and outcomes are colour-coded. This exercise highlights the City's focus on medium-term outcomes like behaviour change and customer satisfaction with comparatively little direct output indicators for the "communication" pathway. It also shows that, although city officials recognize long-term impacts as important, the paucity of indicators identified will likely fail to capture them. Similarly, very few indicators are available to assess the influence of external factors on the project results. The exercise of colour-coding to identify

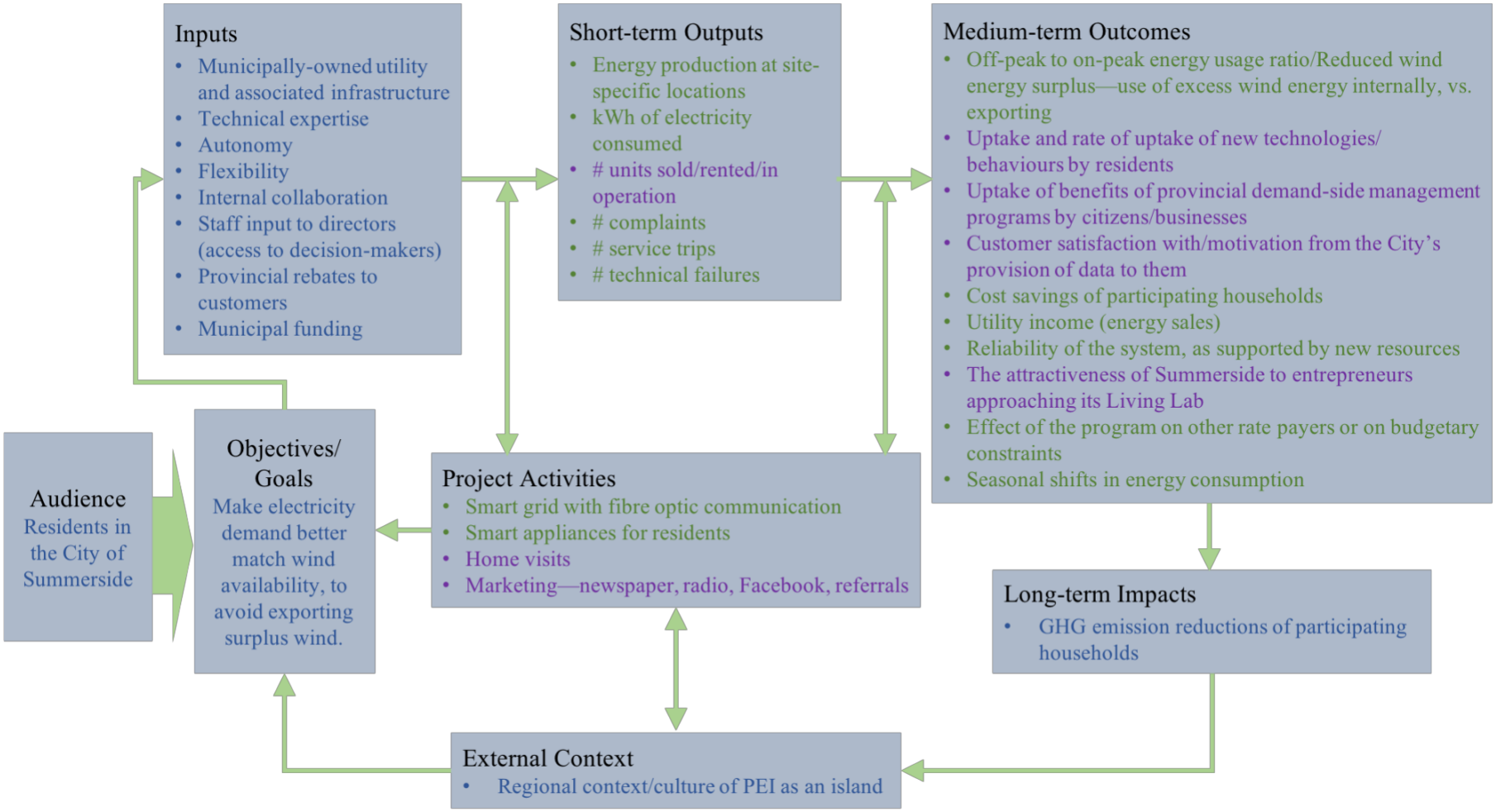

Figure 3. Summerside smart grid theory-of-change model and causal mechanism indicators colour-coded.

the causal mechanisms linking activities to outputs and outcomes has ambiguous results, as the pathways of change are not direct or immediately transparent. A solution is to develop two independent theory-of-change models corresponding to the different causal mechanisms, allowing a deeper understanding of the factors at play.[78] Since a theory of change can be applied at any project level, it is possible to do so.

Working with Summerside, we also illustrate a critical element of theory-driven evaluation—namely, the idea of alternative pathways (Box 3). Here, the aim was to examine if the activities proposed for the project—its implementation characteristics—are the best way to achieve the long-term objectives.[79] The simplified Summerside case illustrates how the chosen path for a transition learning project can be compared to an alternative trajectory and baseline. By exploring *a priori* multiple possible pathways (or counterfactuals), activities where additional gains could be made can be identified. Not unlike the development of a narrative theory of change, examination of alternative pathways may ensure that the project activities have the greatest chances of success.

6.2. The Communauto Car-Sharing Company

Communauto, the oldest car-sharing operator in North America, has provided a station-based car-sharing service to its members since 1994. The service allows members to rent a car from its extensive network of stations for short- or medium-term periods. Users reserve the vehicle ahead of time and return the rented vehicle to its original station. Members are billed according to the rental time and distance travelled. They can opt into several monthly packages depending on their needs and usage. Communauto defines itself as environmentally and socially conscious—its goal is to reduce the number of private vehicles in circulation and to offer a practical and yet affordable solution to access a vehicle.[80] A privately-owned organization, Communauto partners with multiple mobility providers, such as public transit agencies, bike-sharing or train providers, to offer advantageous bundles to members. Besides its markets in Montréal and Québec, Communauto has expanded its operations to multiple provinces in Canada and in France. In this section, SCD scholars who have worked with Communauto for a number of years outline the car-sharing company's evaluations to date.

Communauto's interest in evaluation has been specifically geared towards understanding potential markets and how to grow as a company. A number of studies have looked at the conditions allowing to predict success of car-sharing companies. Parking pressure, vehicle ownership, use of alternative modes of transport and density are predominantly proposed as reasons for the desirability of car-sharing in an area. [81] While it can be used in traditional urban neighbourhood settings, car-sharing can also be available for business use (via company fleets), on university campuses, in large apartment complexes (via partnerships with a promoter), in smaller cities or in rural areas. A large share of car use for commuting trips confirms the high dependency of daily travels on private cars and is not correlated with adoption of car-sharing. Availability of other efficient travel options such as transit or active modes (through walkable neighbourhoods) is, on the contrary, favourable to car-sharing services. While car-sharing is designed to be used for specific needs, a developed mobility cocktail must exist for members to make up for the lack of a vehicle in day-to-day trips. Vehicle ownership is also looked at as car-sharing members mostly have access to one car or none in their household. Density is an important feature of a neighbourhood as a potential pool of new members, but also because it is correlated with other features such as degree of transit service or residential/commercial land-use mix. The proportion of one-person households seemed correlated in the literature with car-sharing success, but it should not be an exclusion factor, considering multi-person households also

Box 3. Technical Illustration of Baselines and Alternative Pathways

The low-carbon energy outcomes of Summerside's smart grid can be evaluated by comparing to the City of Charlottetown. Charlottetown is the other city in Prince Edward Island and is served by Maritime Electric, a regulated public utility that supplies most of the Island. Wind energy makes up roughly 25% of the Maritime Electric supply, and the utility has thus far avoided investing in smart grid infrastructure. While the City of Charlottetown recently completed a GHG inventory and is currently exploring ways to reduce citywide emissions, its options are limited by lack of control over electricity supply and pricing structure. Although local, distributed renewable energy sources such as solar panels are possible, the City has little means of encouraging such developments since electricity prices and net metering policies are set by Maritime Electric. In contrast, Summerside with its own utility can determine its own electricity supply and implement its own pricing structures and demand-response programs, which have allowed it to use double the proportion of wind energy as the provincial baseline. Along with enabling a 50%-wind electricity mix, the Heat for Less Now (HfLN) program is estimated to have reduced annual GHG emissions by 407 tCO_{2e} and to save participants on average $1400/year relative to oil heating.[e]

Although Summerside is one of the most progressive municipalities in Canada for renewable energy integration, it is worthwhile to consider alternative pathways that could improve outcomes further. An alternative model that Summerside could have adopted is the use of real-time electricity pricing. Rather than the existing approach of adjusting heating loads according to a set daily schedule, real-time pricing would set electricity prices dynamically based on the current availability of renewable energy and the level of demand in the grid. Such an approach offers advantages in terms of both integrating renewables and raising consumer awareness. One of the goals of Summerside Electric is to maximize internal use of its wind power and to minimize periods when there is excess wind energy that must be exported—since these intermittent exports typically receive unfavourable prices. This corresponds with maximizing the return on the wind farm investment and reducing reliance on external electricity grids to compensate for renewable energy fluctuations. Summerside Electric estimates that its existing HfLN smart grid program provided a 17% reduction in excess wind power in 2015.[e] Applying a simple supply and demand analysis method[f] to hourly electricity data for Summerside for 2017 shows the difference real-time pricing could make. With the assumption that a real-time pricing approach could leverage similar demand response capacities as the existing HfLN program, the analysis suggests that the adaptability of real-time pricing results in three to four times greater reductions in excess wind energy than the existing scheduled approach. Beyond its greater potential effectiveness in increasing use of wind energy, a real-time pricing approach also has the potential to induce a shift in consumer thinking. Whereas static pricing does not provoke any awareness of electricity consumption timing, and time-of-use pricing raises awareness of typical times of peak, real-time-pricing reminds the consumer that electricity supply depends on the availability of the wind on an ongoing basis, giving a financial reason to be conscious of where the energy is coming from.

join car-sharing programs. Finally, other intangible factors such as the implication of a committed spokesperson, support from the City (such as access to parking spaces for shared cars), funds, promotion or car-sharing education are one of the many other considerations for the success of a car-sharing scheme.

Communauto has partnered with researchers to evaluate the environmental impacts of its shared fleet of vehicles. In one such study,[82] researchers asked respondents to provide information on their estimated car mileage and ownership before and after joining Communauto. Car mileage is used as

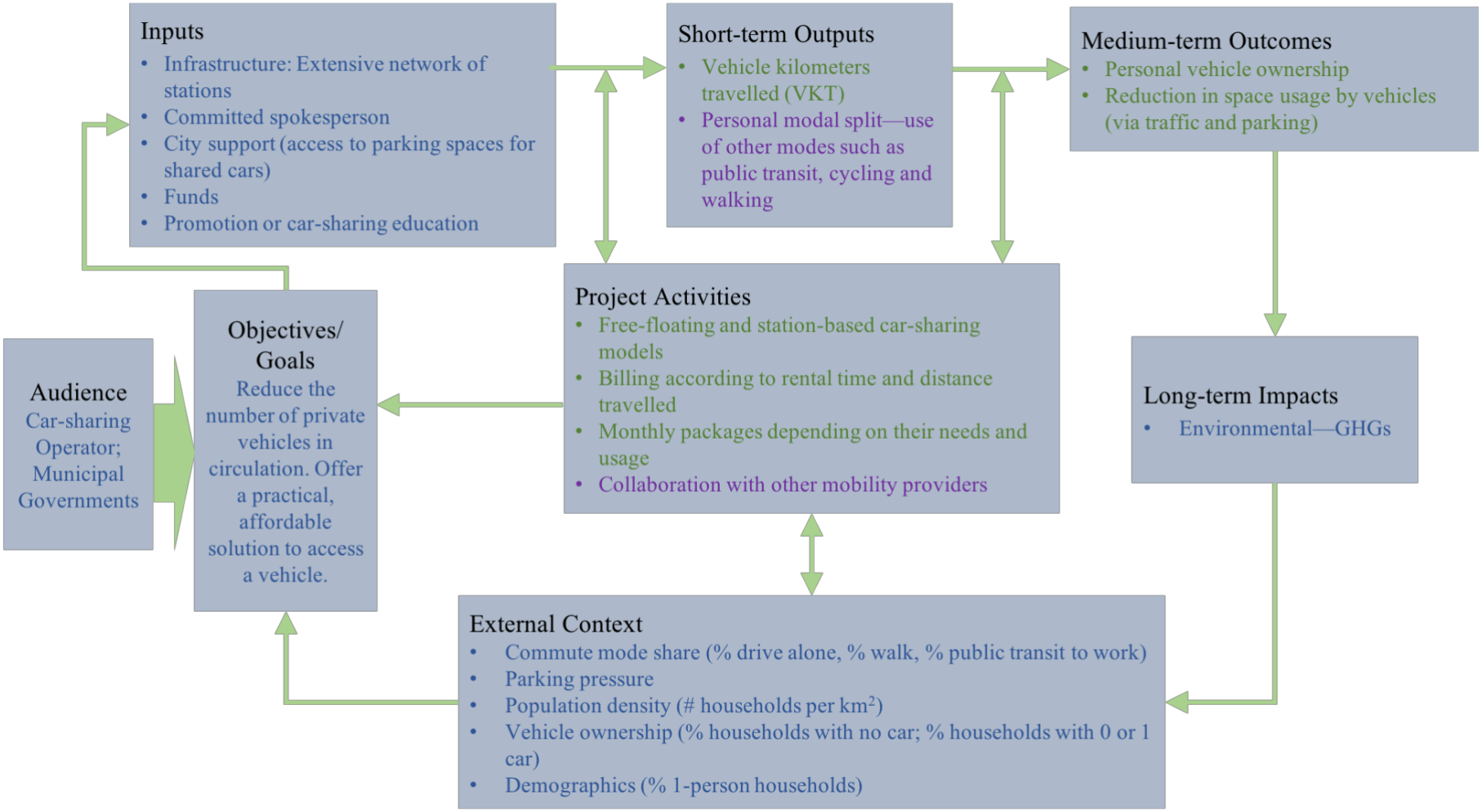

Figure 4. Communauto theory-of-change model and causal mechanism indicators colour-coded.

an indicator to evaluate the difference in vehicle kilometers travelled, which could be translated into CO_2 reduction, a metric widely used by the scientific community and decision-makers. Vehicle ownership represents the potential of vehicle reduction, as one benefit of car-sharing can be the reduction in space usage by vehicles via traffic and parking. In another study on car-sharing members' mobility behaviour,[83] a specific origin-destination travel survey for car-sharing members was performed in fall 2008, in parallel to one conducted by the regional transit agency. This led to a comparison between travel behaviour of car-sharing members and the general population. Members had to describe all performed activities, their corresponding locations and the mode of transport used for one recent weekday. We derived the following narrative theory of change from these two studies:

Context. A car-sharing company, Communauto, aims to reduce the number of private cars on the road while providing an affordable alternative.

Causal mechanisms. (1) Access pathway: Making station-based fleets of shared vehicles accessible allows members to easily opt in to car-sharing over private-car trips. (2) Flexibility pathway: Collaborating with other mobility providers facilitates users' use of shared versus private cars.

Figure 4 exemplifies how existing indicators could be used to test the causal mechanisms at the heart of theory-driven evaluation, thus illustrating the flexibility of the proposed evaluation framework. Unlike Summerside, for which the indicators selected by city officials allowed close examination of middle-term outcomes, the indicators used to date by Communauto largely focus on immediate outputs, leaving relatively little knowledge on potential pathways for change. This is not surprising given Communauto's own motivation for evaluation and shows the importance of a federal-government-led program of transition learning for evaluation to inform a large-scale low-carbon transition.

6.3. Toronto and Region Conservation Authority's Sustainable Neighbourhood Retrofit Action Plan

The Sustainable Neighbourhood Retrofit Action Plan (SNAP) is a collaborative, neighbourhood-based solution for advancing urban renewal and climate action in older urban areas. SNAP helps municipalities improve energy efficiencies, draw strong local community support and build innovative partnerships for implementation of a broad range of initiatives in the public and private realms. By reframing environmental projects to incorporate greater social and economic outcomes, SNAPs are intended to help generate creative solutions that garner more support for implementation. Most importantly, SNAPs deliver measurable results and scalable lessons toward the implementation of watershed plans and other municipal plans and strategies (e.g., sustainable development, climate action and economic and social development). Individual SNAP Action Plans have four typical actions areas: (1) home retrofit programs that increase uptake in multiple sustainability, quality of life and resilience actions, including climate actions at home, and also encourage community life beyond the private home; (2) integrated infrastructure and public realm renewal projects; (3) integrated site projects and demonstrations on Multi-unit residential and Industrial Commercial and Institutional buildings; and (4) increasing community resilience and community connections.

To understand their visions for a useful framework to evaluate the SNAP projects, discussions were held with five Toronto and Region Conservation Authority (TRCA) employees at the TRCA offices in July 2018 (see full meeting summary in Appendix 3). According to the TRCA, the value of evaluation is both external and internal. Externally, it serves for reporting to funders on outcomes and to potential partners that the TRCA can attract and retain. Internally, it stimulates continuous improvement of the program to be more cost-efficient and assess whether a new approach is needed for reaching impacts, which elements that work to expand, proximity to targets and where to shift or adapt. Evaluation can confirm whether the design for a neighbourhood typology is working and inform scaling the project to other similar neighbourhoods in the Toronto Region and across Canada.

For its single family residential retrofit programs, the SNAP team had developed their own evaluation using a linear logic model approach that follows four high-level methods for comparison: within neighbourhoods; between neighbourhoods; with mass marketing programs; and with control groups. The new West Bolton SNAP, for example, is set to undergo an Interim Evaluation to assess change in behaviour, attitudes and retrofit uptake as a result of the pilot Home Retrofit Program (HRP) (early 2019), to evaluate the effectiveness of and identify ways to improve the Residential HRP, and to inform similar programs in other SNAP neighbourhoods. Additional evaluation activities are expected to occur after the Interim Evaluation by the SNAP team to assess long-term outcomes of the project—for example, estimating GHG savings based on energy retrofit technologies or behaviours adopted.

In terms of climate change mitigation and water conservation, the TRCA's original idea was to evaluate progress based on changes in actual neighbourhood-scale water and energy use and GHG emissions. That approach was eventually deemed unworkable due to constraints in data availability from utilities related to privacy concerns. The TRCA then decided to evaluate behavioural change and the uptake of energy retrofit actions instead. For most SNAPs, the TRCA is undertaking evaluation in-house. For SNAP's newest West Bolton single family residential retrofit program, however, the TRCA opted to hire a consultant to deliver the evaluation workplan on a condensed timeline in response to these challenges. They see the benefit of a consultant in offering a third-party perspective.

Based on the information gathered, we developed the following narrative theory of change for SNAP's residential retrofit programs:

Context. Mass-marketed residential home retrofit incentive programs (e.g., water and waste conservation rebates) are not resulting in uptake levels required to achieve sustainability objectives (e.g., GHG reduction targets). The TRCA's goal is to build community-level awareness and engagement with regional sustainability issues, including climate change.

Causal mechanisms. (1) Community pathway: Neighbourhood-specific, locally-targeted residential engagement through neighbourhood events, workshops, one-on-one home consultations, and other tactics can help build community awareness and empower residents to take actions, including accessing provincial and utility rebates for home retrofit investments. This high-touch community-scale engagement, in addition to mass marketed retrofit programs, can create the conditions for individual behaviour change as well as spark broader systemic change in support of sustainability transitions. (2) Homeowner pathway: Targeting individuals and supporting homeowners through face-to-face consultations and ongoing support and follow-up will increase uptake and continued participation. (3) Partnership pathway: Working in close partnership with existing municipal programs and incentives will catalyse engagement and cross-promote participation.

The proposed evaluation framework and indicators based on the theory-of-change model articulated above (Figure 5) bear a few key differences with SNAP's own logic model. The TRCA has been very diligent at identifying a large range of indicators that could serve to monitor outputs, outcomes and impacts. Our proposed evaluation framework seeks to illustrate multiple links between the building blocks of the theory-of-change model to emphasize how they can influence each other. Being linear, the logic model proposed by the SNAP team cannot capture some of those complex links, making it difficult in turn to determine the causes of project success or failure. It has been suggested that too

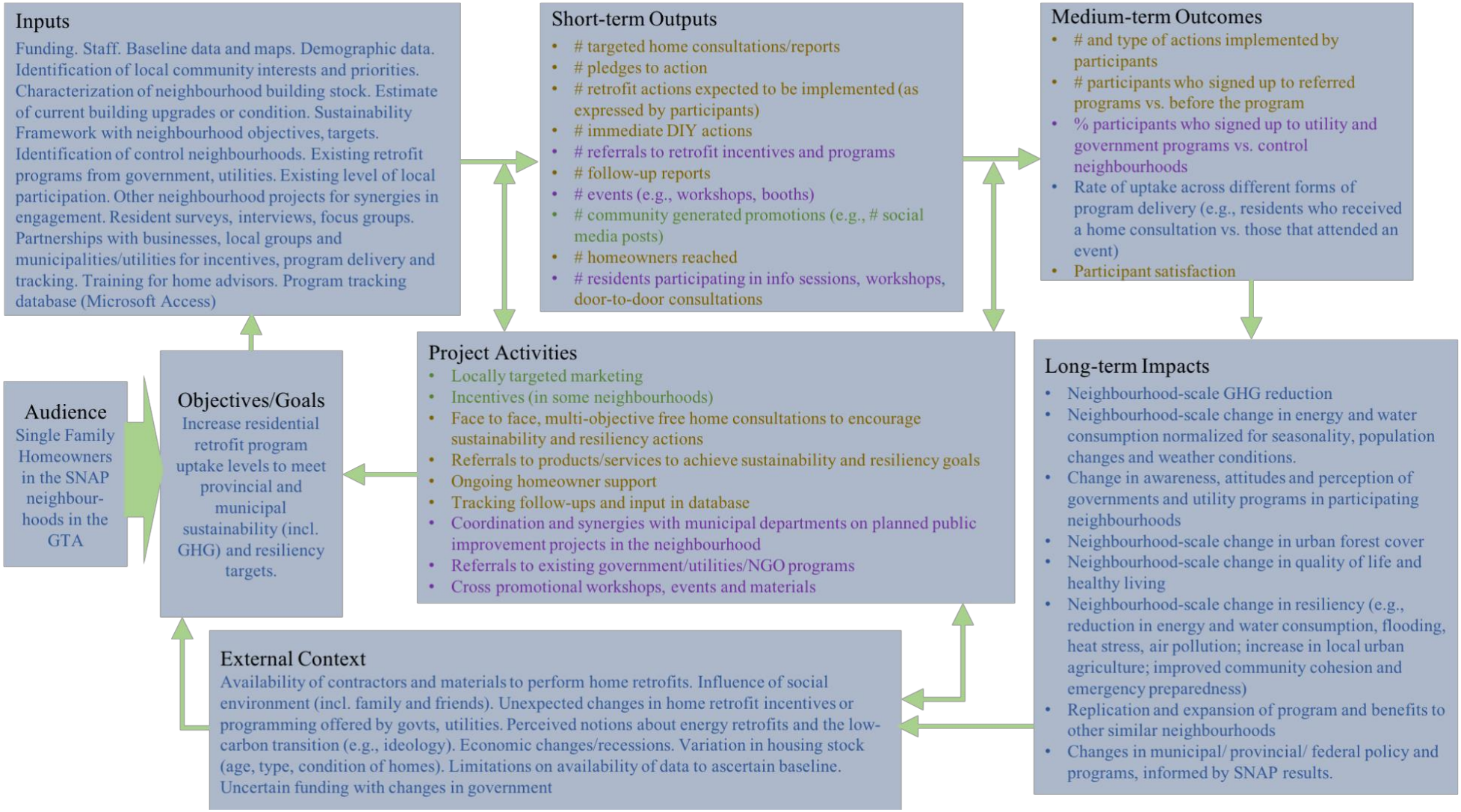

Figure 5. SNAP single family residential retrofit program theory-of-change model and causal mechanism indicators colour-coded.

much detail in the development of the logic model and the resulting evaluation could become inefficient by both blurring the causal relationship and causing evaluation exhaustion.[84] As it strives to evaluate SNAP, the TRCA identified a number of challenges many of which pertained to financing, skillsets and capacity. We suggest that if evaluation simply flows from the initial definition of a solid theory-of-change model, using a reduced but highly pertinent number of indicators, the burden of evaluation would be lessened, and its usefulness increased.

7. THE EVALUATION PROCESS

After developing a robust theory-of-change model, it is crucial to realize that *how* to evaluate cannot be disconnected from the context of the project being implemented. Instead, there are important practical questions to ask in developing a project-specific evaluation framework, precisely because different answers to these questions will produce different needs for and kinds of appropriate evaluation procedures and criteria.

For whom is the evaluation intended?

The evaluation framework proposed explicitly considers the audience of the project as a structural element of output, outcomes and impacts (Figure 2). The question "For whom is the evaluation intended?" might, however, have different answers. The evaluation's users will not necessarily be the audience of the project. For example, for both the TRCA and Summerside, it is clear that their audience is the local residents. However, the TRCA considers that evaluation is intended for a broad range of stakeholders, including funders, municipal, provincial and federal government, SNAP program managers and the community. In Summerside, an evaluation would be intended for both internal and external partners: city council; federal/provincial agencies; city administrators, senior bureaucrats and staffers in economic development, at the electric utility and in recreation/leisure, and the entrepreneurs who approach the City for collaboration.

This question forces us to recognize that there are a number of participants and stakeholders in low-carbon transitions with diverse interests in evaluation. It is unlikely that a single evaluation procedure or set of criteria will satisfy all of them. Thus, either multiple evaluation procedures need to be employed or the audiences for evaluation need to be prioritized. Let's do a thought experiment with a Sustainable Neighborhood Retrofit Action Plan: Who might want to know if the experiment is working and why do they need to know this/what might they do with this information? The answers to these questions might complement one another (e.g., where some evaluation information can be used by multiple audiences), other times they will not, and sometimes they will be nested.

- *Households participating* in retrofitting might want to understand if their life has gotten better in some tangible way, whether participating was a good return on their investment (assuming that households need to provide some capital), how many of their neighbours have participated in the program and the impact that the program is having in the overall neighbourhood. This will have implications for community pride, but also for the value of houses in the neighbourhood. After going through the program, homeowners will have a solid understanding of the exponential impact of collective action—e.g., if everyone manages stormwater sustainably, flooding in the overall neighbourhood will be reduced; if everyone plants a tree, heat island and air quality will improve.
- *Providers of retrofitting technology* will want to know ways to increase uptake. They are also interested in efficiencies associated with the neighbourhood scale (i.e., being able to implement the technology in a geographical area of manageable size)
- *The organization* implementing or running the project (e.g., an NGO or neighbourhood association) will want to know whether the program works (i.e., has the desired outcomes) at a reasonable cost, if the program needs to be adjusted as the project is running or preparing for future phases of the project and what contribution the project is making to the transition.

- *The municipal government* will want to know if the project is suitable for city-wide rollout if considered successful.
- *The provincial government* will want to know if the project is appropriate for other cities in the province.
- *The national government* will want to know if the project is appropriate for cities countrywide.
- *The Transnational City Network and/or other projects' proponents* might want to learn about what works in this project to alter some of their own experimental activities.

In every project, then, there will be two kinds of broad interests in evaluation, though there will be overlap: (1) Some actors are most concerned with the specific performance of the project activities for their own sake (as in the Communauto case study). They are concerned with the *in situ* substantive performance of the project to make it work better for themselves or for their organization. For them, the project may be an end in itself. What counts as "working" will have specific connotations related to the project's performance metrics—e.g., behavioural, physical, monetary, and so on. (2) There will be other actors concerned with the success of the project in those performance metrics *and* its ability to either replicate (e.g., trying the experiment in other neighbourhoods) or scale up (e.g., expanding from the neighbourhood to the whole city). Evaluation will likely include contextual factors and political considerations.

Who is carrying out the evaluation?

There are two models or ideal types of evaluator: (1) *The classic external actor*. Academics or consultants bring an evaluation procedure to the project and conduct the evaluation, which can be more or less participatory. (2) *The co-production model*. Project proponents, especially those doing the experimental activity and perhaps broader stakeholders, collaborate on the evaluation procedure and the treatment or research design that will dictate which conclusions can be drawn and the statistical analyses. The latter model was emphasized during the June 1 workshop. Academics or consultants could participate in roles ranging from partners in carrying out evaluation to serving in a more limited advisory capacity.

For the TRCA, involving a third party in evaluation is important for the evaluation to be seen as legitimate and valid by prospective funders or other partners. Academic partners lend more credibility and could also provide the means to verify a theory-of-change model developed in a participatory way against the existing evidence. Summerside considers that a third-party approach removes bias—most people would not veer from a path once they have committed to it. Since the City does not have an evaluation department, evaluation could be a distributed, delegated function of perhaps any department that can use it. If the intent of evaluation is to sell more product, the economy or industry expects third-party evaluation or standards.

How are the results shared?

This question is related to the purpose of the evaluation. Evaluation that is geared towards improving the performance of a specific ongoing project will be shared differently than a *post-hoc* evaluation of whether a project was "worth it" and ripe for scaling-up and replication. Given that specific projects rarely have the mission of scaling-up the low-carbon transition, SCD scholars have proposed the establishment and coordination of a **pan-Canadian Transition Learning Program** that facilitates learning around the multiple dimensions of scaling-up (Appendix 4). A **Transition Learning Program**, as opposed to a **Transition Learning Project**, would put sharing at the very heart of its activity and give a larger weight to indicators of long-term impacts rather than immediate outputs.

For the TRCA, year-end progress reports to municipal partners to support budget requests, and final reports for grants are a logical way to share results; local, provincial, national and international

conferences and webinars and award applications are additional means. Results are also shared verbally and informally with community members as a tactic to increase engagement. As part of the original provincial funding, the idea was to create a knowledge-sharing network or 'community of practice' to share lessons among behaviour-focused interventions that received money from the same Partners for Climate Action grant. (However, that funding was pulled.) For SNAP, the core interest as it relates to evaluation is in continuous improvement of their own program and providing a basis for communicating impact to funders and other stakeholders.

In Summerside, evaluation results could be shared via internet or electronic reports. If they were fit for public consumption, the City would make the results available on its website, and they would be shared with people like SCD. The City has shared much with its peers—other municipalities, associations, engineering groups and professional affiliates. The City would export lessons to others if triggered by a request; it would not be self-directed. The City would share evaluations if they were available with federal or provincial funding partners, if it were making an application for any type of energy-related project. The results of evaluation need narrative explanations and visuals, not just data; they would have to be digestible, intelligible and utilitarian to be useful.

Who funds the evaluation?

Funding evaluation was mentioned by both the TRCA and Summerside as a significant impediment and was highlighted in the June 1 workshop. The TRCA uses its own in-house staff time, primarily through municipal funding, and senior government grant funding (e.g., Federation of Canadian Municipalities). June 1 workshop participants emphasized the shortcomings of an approach that "diverts money from the project to evaluation, thus making it more difficult to reach the stated goals". In Summerside, federal assistance would be needed to contribute to a national agenda since the City's primary function is to provide services and not to do evaluative research. We suggest that if learning is an important element of project design, and given the proposal to build **Transition Learning Projects** around their own theory-of-change models, the evaluation costs becomes an important element of project implementation itself rather than an external cost (see Box 4 for an example of international leadership).

Box 4. Sweden's National Innovation Arena, RE:Source

The Swedish Energy Agency has launched RE:Source, a 'strategic innovation program' intended to bring together industry, resource and waste management and research sectors to make Sweden a leader in the transition to a circular economy.[g] Innovations include novel technologies and business models for sustainable resource and waste management, for which innovators can also get program funding. Objectives include sharing lessons within the program and with the general public to "influence business offerings, international standards, ecolabelling, policies and the general population's awareness"; education and creation of industry expertise; international export of Swedish products and services; and new jobs.[g] The ultimate goals are to improve resource efficiency and competitiveness and address major societal challenges like climate change and energy supply.[h]

8. CONCLUSION

This report focusses on the evaluation of **Transition Learning Projects** that could serve as important elements in the acceleration of the low-carbon transition. The very characteristics of such **Transition Learning Projects**, as summarized in Box 5 below, are therefore fundamental components of the development of an evaluation framework.

Box 5. Transition Learning Projects should:

- Innovate and link with overarching goals and priorities of low-carbon energy transition pathways
- Be framed as learning projects—project proponents must articulate a clear goal to learn from project experiences
- Balance and integrate technological, business and social innovation
- Involve collaboration across multiple societal sectors (businesses, NGOs and governments)
- Have a public engagement component
- Provide opportunities for structured assessment and learning.

Taken together they should:

- Cover different dimensions of the energy transition (housing, transport, electricity production and consumption) in different regions and cities, business sectors, and so on
- Be structured around a set of strategic priorities that takes account not just of potential short-term gains (more readily accessible emission reductions) but overcoming key bottlenecks and opening the door to further long-term change
- Mobilize researchers and technical experts to assist projects and systematize learnings
- Participate in systematic evaluation of projects and lessons across experiments, while making data (quantitative metrics and qualitative lessons) widely accessible to encourage further knowledge creation
- Have a substantial outreach dimension to engage and educate Canadians.

We propose that **Transition Learning Projects** would be best evaluated by adopting a theory-of-change framework. We further suggest that the evaluation should focus not only on the inputs, outputs and outcomes but also on the causal relationships, the effectiveness of the process and the long-term impacts including on policymaking, on society broadly and on the low-carbon transition. To fully capture the complexity of social systems, the framework must be non-linear, explicitly accounting for contextual factors that will likely influence them. In addition, the evaluation approach must be adaptive in response to inevitably changing system conditions. The complexity of sustainability transitions underlines the importance of embracing transdisciplinary (science and social science) and multi-method (quantitative and qualitative) approaches.

Theory-driven evaluation, as we propose, can be carried out in discrete steps as follows (summarized from [85]):

1. Determine the final objective of a project in terms of the low-carbon transition, ideally in a participatory way including key stakeholders.
2. With the key stakeholders, develop a theory-of-change model linking the final goal with possible steps to get there, taking into account causal relationships. This could be done in a half-day workshop, for example.
3. Verify the causal relationships with published evidence.
4. Develop a narrative for the proposed theory of change as a way to validate it by clearly articulating the assumptions on which the theory-of-change model is built.
5. Validate the "final" theory-of-change model with the key stakeholders.
6. Identify the implementation characteristics—i.e., the activities that the project will undertake to meet its final objective—while paying attention to alternative pathways, ensuring that the

activities are those best-suited to meet the final project objective and test the causal relationships identified in the theory-of-change model.

7. Identify indicators for all theory-of-change model components—inputs, outputs, outcomes and impacts.

An important advantage of the theory-driven evaluation that we propose is that it could serve to identify whether project failure is due to implementation failure, engagement failure or theory-of-change failure. Four questions can be asked at the end of the evaluation to provide such insight: (1) Was the project adequately implemented? (2) Was there sufficient uptake, engagement and adherence? (3) Were intermediate outcomes achieved? (4) Were final outcomes achieved?[86] The ability to identify why a project fails is an essential element of de-risking low-carbon energy projects, as identified by participants in the June 1 workshop. It also opens the door for beginning to scale up the lessons learned and accelerate the low-carbon transition.

In this report, we have illustrated the proposed evaluation framework, examining three case studies. This exercise helped us identify some aspects of evaluation that would be modified if the proposed evaluation framework were to be adopted. For example, we noted an apparent lack of initial review of precedents and literature or research findings to determine if the relevant interventions have proved effective elsewhere. Discussions around evaluation with practitioners suggested that they tended to focus primarily on finding indicators rather than first identifying pathways for change, which is of fundamental importance if the lessons from **Transition Learning Projects** are to be scaled up. Likewise, the case studies also indicate that longer-term, intangible, societal impacts of the project are given less attention in practitioners' contemplations of evaluation.

While recognizing that we were tasked with developing a framework to evaluate socio-technical energy **Transition Learning Projects**, we contend that more could be learned by considering lessons learned from an *array* of such projects implemented to accelerate and scale up the low-carbon transition in Canada. Such a holistic approach would open and energize a space from which lessons learned can help promote deployment of appropriate and effective low-carbon alternatives, shift decisions made by consumers and develop evidence-based policy (see Box 6 for an international example).

Box 6. Energy Transition in the Netherlands as a model to elicit long-term impacts

In 2002, the Dutch Ministry of Economic Affairs initiated a government transition management program in the Netherlands. Seven 'platforms' were created: chain efficiency (production and consumption); bio-based raw materials; new gas; sustainable mobility; sustainable electricity supply; built environment; and greenhouses as an energy source. [i,j] Public, private and community sector representatives from the platforms defined their visions, set their transition paths and objectives and selected 'frontrunners' to establish their own experiments.[i] The government set up the 'frontrunner desk' to spur innovation by mediating between business and government to identify policy changes needed to remove barriers to innovation.[k] Where innovators required financing, access to decision-makers, support to navigate legalities and permits, and help in networking and marketing, for example, financial instruments were adjusted, cooperation among and within government ministries encouraged and policy developed.[k] During its operation (2004–2006), the frontrunner desk solved more than 59% of the challenges confronted by the 69 companies that approached it for help navigating barriers to innovation.[l] To promote learning-by-doing, transition projects and paths were subjected to ongoing monitoring, analysis and evaluation by an overarching special council.[l]

Indeed, lessons about causal relationships as tested by theory-driven evaluations are likely to only apply to precise systems (e.g., home retrofit, farms, mobility, etc.) and specific contexts (e.g., provinces, urban, etc.). Thus, only evaluation of a large number of **Transition Learning Projects** could allow identification of fundamental elements of a general theory of change for the low-carbon transition (e.g., existence and availability of alternative technologies, attitudes or changes in attitudes of groups of peers, and so on). We therefore propose the creation by the Government of Canada of a low-carbon energy, **pan-Canadian Transition Learning Program** (see scholars' detailed proposal in Appendix 4). It would be a showcase to the world of Canada's socio-technological approach to promoting a low-carbon energy transition.

9. REFERENCES

References in Boxes

[a] OECD. 2017a. *Embracing innovation in government: Global trends 2017*. Paris: OECD Publications.
[b] OECD. 2017a. *Embracing innovation in government: Global trends 2017*. Paris: OECD Publications. P. 54.
[c] Matschoss, K., and Heiskanen, E. 2017. Making it experimental in several ways: The work of intermediaries in raising the ambition level in local climate initiatives." Journal of Cleaner Production.
[d] https://www.ofgem.gov.uk/about-us/how-we-engage/innovation-link
[e] Wong, S., G. Gaudet, and L.-P. Proulx. 2017. Capturing Wind with Thermal Energy Storage—Summerside's Smart Grid Approach. IEEE Power and Energy Technology Systems Journal, 4(4): 115–124.
[f] Hall, M. and A. Swingler. 2017. Initial Perspective on a 100% Renewable Electricity Supply for Prince Edward Island. International Journal of Environmental Studies, 75(1): 135–153.
[g] https://www.vinnova.se/en/m/strategic-innovation-programmes/resource/
[h] http://www.energimyndigheten.se/forskning-och-innovation/forskning/bioenergi/bransleforsorjning/program/ resource/
[i] Veilleux, J.-P. 2010. Energy Transition in the Netherlands. PRI Project Sustainable Development. Government of Canada.
[j] Kern, F. and Smith, A. 2008. Restructuring energy systems for sustainability? Energy transition policy in the Netherlands. Energy Policy, 36(11): 4093–4103.
[k] Kemp, R. 2010. The Dutch energy transition approach. International Economics and Economic Policy, 7: 291–316.
[l] Mytelka, L. et al. 2012. Policies for capacity development. In N. Nakicenovic (Ed.) Global energy assessment: Towards a sustainable future. International Institute for Applied Systems Analysis.

References in Text

[1] Geels, F.W., and J. Schot. 2007. Typology of sociotechnical transition pathways. Research Policy, 36(3): 399–417.
[2] Smith, A., A. Stirling, and F. Berkhout. 2005. The governance of sustainable socio-technical transitions. Research Policy ,34(10): 1491–1510.
[3] Senbel, M., V.D. Ngo, and E. Blair. 2014. Social mobilization of climate change: University students conserving energy through multiple pathways for peer engagement. Journal of Environmental Psychology, 38: 84–93.
[4] Nilsson, M., L. J. Nilsson, R. Hildingsson, J. Stripple, and P.O. Eikeland. 2011. The missing link: bringing institutions and politics into energy future studies. Futures, 43(10): 1117–1128.
[5] Loorbach, D., and J. Rotmans. 2010. The practice of transition management: Examples and lessons from four distinct cases. Futures ,42(3): 237–246.
[6] Gross, M. 2010. Ignorance and Surprise: Science, Society, and Ecological Design. Cambridge, MA: The MIT Press.
[7] Rosenbloom, D., J. Meadowcroft, S. Sheppard, S. Burch, and S. Williams. Transition experiments: Opening up low-carbon transition pathways for Canada through innovation and learning (submitted). The discussion paper for Sustainable Canada Dialogues' side-event at Generation Energy in October 2017, which is based on Rosenbloom et al. submitted, is available at: https://www.researchgate.net/publication/320161503.
[8] Bernstein, S., and M. Hoffmann. 2018. The politics of decarbonization and the catalytic impact of subnational climate experiments. Policy Sciences, 51(2): 189–211.
[9] Kivimaa, P., M. Hildén, D. Huitema, A. Jordan, and J. Newig. 2017. Experiments in climate governance – A systematic review of research on energy and built environment transitions. Journal of Cleaner Production, 169: 17–29.
[10] Sengers, F., A. J. Wieczorek, and R. Raven. 2016. Experimenting for sustainability transitions: A systematic literature review. Technological Forecasting and Social Change.
[11] van den Bosch, S., and J. Rotmans. 2008. Deepening, broadening, and scaling up: a framework for steering Transition Experiments. Rotterdam: Knowledge Centre for Sustainable System Innovations and Transitions.
[12] Brown, H. S., and P.J. Vergragt. 2008. Bounded socio-technical experiments as agents of systemic change: The case of a zero-energy residential building. Technological Forecasting and Social Change, 75(1): 107–130.
[13] Romijn, H., R. Raven, and I. de Visser. 2010. Biomass energy experiments in rural India: insights from learning-based development approaches and lessons for strategic niche management. Environmental Science & Policy, 13(4): 326–338.
[14] Bulkeley, H., V. Castán-Broto, and G. Edwards. 2015. An urban politics of climate change: experimentation and the governing of socio-technical transitions. London; New York: Routledge, Taylor & Francis Group.
[15] Rosenbloom, D., J. Meadowcroft, S. Sheppard, S. Burch, and S. Williams. Transition experiments: Opening up low-carbon transition pathways for Canada through innovation and learning (submitted). The discussion paper for Sustainable Canada Dialogues' side-event at Generation Energy in October 2017, which is based on Rosenbloom et al. submitted, is available at: https://www.researchgate.net/publication/320161503.

[16] Luederitz, C., N. Schäpke, A. Wiek, D. J. Lang, M. Bergmann, J. J. Bos, S. Burch, … F. R. Westley. 2016. Learning through evaluation – A tentative evaluative scheme for sustainability transition experiments. Journal of Cleaner Production, 169: 61–76.
[17] van der Ven, H., S. Bernstein, and M. Hoffmann. 2017. Valuing the contributions of nonstate and subnational actors to climate governance. Global Environmental Politics, 17(1): 1–20.
[18] http://energyfutureslab.com
[19] Williams, S. (Forthcoming). The Alberta Energy Futures Lab: A Case Study in Socio-Cultural Transition through Public Engagement in Creating Spaces of Engagement. Eds. Wiebe, S. & Ney, T. University of Toronto Press.
[20] Park, T. 2017. Evaluating the Nest Learning Thermostat Four field experiments evaluating the energy saving potential of Nest's Smart Heating Control. Behavioural Insights Team..
[21] Pearl and Mackenzie. The Book of Why: The New Science of Cause and Effect. Working Draft.
[22] Arnold, E., T. Åström, C. Glass, and M. De Scalzi. 2018. How should we evaluate complex programmes for innovation and socio-technical transitions? Brighton, UK: Technopolis Group.,
[23] Rotmans, J., and D. Loorbach. 2009. Complexity and Transition Management. Journal of Industrial Ecology, 13(2): 184–196.
[24] Andersson, C. 2014. Complexity science and sustainability transitions. Environmental Innovation and Societal Transitions, 11: 50–53.
[25] Abbott, K.W. 2017. Orchestrating experimentation in non-state environmental commitments. Environmental Politics, 26(4): 738–763.
[26] The literature review builds on Williams, S. (2017, June). Evaluating Transition Experiments in Times of Rapid Change. Paper presented at International Sustainability Transitions 2017. Chalmers University, Gothenburg, Sweden. For additional literature, see Appendix 4 and Excel document, Evaluation Framework Literature Review.
[27] See, for example, Raven et al. 2017. Unpacking sustainabilities in diverse transition contexts: solar photovoltaic and urban mobility experiments in India and Thailand. Sustainability Science, 12: 579–596, p. 64; and Kivimaa, P., M. Hildén, D. Huitema, A. Jordan, and J. Newig. 2017. Experiments in climate governance – A systematic review of research on energy and built environment transitions. Journal of Cleaner Production, 169: 17–29.
[28] Westley, F., P. Olsson, C. Folke, T. Homer-Dixon, H. Vredenburg, D. Loorbach, et al. 2011. Tipping Toward Sustainability: Emerging Pathways of Transformation. Ambio, 40(7): 762–780.
[29] Rotmans, J., and D. Loorbach. 2009. Complexity and Transition Management. Journal of Industrial Ecology, 13(2): 184–196.
[30] Yukawa, J. 2015. Preparing for Complexity and Wicked Problems through Transformational Learning Approaches. Journal of Education for Library and Information Science Online, 56(2): 1–11.
[31] Arnold, E., T. Åström, C. Glass, and M. De Scalzi. 2018. How should we evaluate complex programmes for innovation and socio-technical transitions? Brighton, UK: Technopolis Group.
[32] Rotmans, J., and D. Loorbach. 2009. Complexity and Transition Management. Journal of Industrial Ecology, 13(2): 184–196. P. 17.
[33] Schot, J., and L. Kanger. 2018. Deep transitions: Emergence, acceleration, stabilization and directionality. Research Policy, 1–15.
[34] Smil, V. 2010. Energy Transitions. Santa Barbara, CA: ABC-CLIO.
[35] Grübler, A. 1998. Technology and Global Change. Cambridge: Cambridge University Press.
[36] Hölscher, K., J.M. Wittmayer, and D. Loorbach. 2017. Transition versus transformation: what's the difference? Environmental Innovation and Societal Transitions, 27: 1–3.
[37] Arnold, E., T. Åström, C. Glass, and M. De Scalzi. 2018. How should we evaluate complex programmes for innovation and socio-technical transitions? Brighton, UK: Technopolis Group.
[38] Rotmans, J., and D. Loorbach. 2009. Complexity and Transition Management. Journal of Industrial Ecology, 13(2): 184–196. P. 39.
[39] Rosenbloom, D., J. Meadowcroft, S. Sheppard, S. Burch, and S. Williams. Transition experiments: Opening up low-carbon transition pathways for Canada through innovation and learning. Canadian Public Policy (in press).
[40] Abelson, J., P.-G. Forest, J. Eyles, P. Smith, E. Martin, and F.-P. Gauvin. 2003. Deliberations about deliberative methods: issues in the design and evaluation of public participation processes. Social Science & Medicine, 57(2), 239–251.
[41] Black, L., S. Burkhalter, J. Gastil, and J. Stromer-Galley. 2008. Methods for Analyzing and Measuring Group Deliberation. In L. Holbert (Ed.), Sourcebook of Political Communication Research Methods, Measures and Analytical Techniques. New York.
[42] Carpini, M., F.L. Cook, and L.R. Jacobs. 2004. Public deliberation, discursive participation, and citizen engagement: A review of the empirical literature. Annual Review of Political Science, 7(1): 315–344.

[43] Barrett, G., M.Wyman, and V.S.P. Coehlo. 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.
[44] Caddy, J. 2005. Evaluating public participation in policy-making. OECD Publishing.
[45] Gaventa, J., and G. Barrett. 2010. So What Difference Does it Make? Mapping the Outcomes of Citizen Engagement. IDS Working Papers, 2010(347), 01–72.
[46] Mutz, D.C. 2008. Is Deliberative Democracy a Falsifiable Theory? Annual Review of Political Science, 11(1): 521–538.
[47] 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.
[48] Ryfe, D.M. 2005. Does Deliberative Democracy Work? Annual Review of Political Science, 8(1): 49–71.
[49] Darcy Riddell, personal communication, April 2017.
[50] Wiek, A., S. Talwar, M. O'Shea, and J. Robinson. 2014. Toward a methodological scheme for capturing societal effects of participatory sustainability research. Research Evaluation, 23(2): 117–132.
[51] Cornish, L. 2013. Can 4D visioning foster community responses on climate change? Master's Thesis, University Of British Columbia.
[52] Robinson, J., and J. Tansey. 2006. Co-production, emergent properties and strong interactive social research: the Georgia Basin Futures Project. Science and Public Policy, 33(2): 151–160.
[53] 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.
[54] Blackstock, K. L., G.J. Kelly, and B.L. Horsey. 2007. Developing and applying a framework to evaluate participatory research for sustainability. Ecological Economics, 60(4) : 726–742.
[55] Currie, M., G. King, P. Rosenbaum, M. Law, M. Kertoy, annd J. Specht. 2005. A model of impacts of research partnerships in health and social services. Evaluation and Program Planning, 28(4): 400–412.
[56] Kasemir, B., J. Jager, C.C. Jaegar, and M.T. Gardner. (Eds.). 2003. Public Participation in Sustainability Science: A Handbook. Cambridge: Cambridge University Press.
[57] Robinson, J., S. Burch, S. Talwar, M. O'Shea, and M. Walsh. 2011. Envisioning sustainability: Recent progress in the use of participatory backcasting approaches for sustainability research. Technological Forecasting & Social Change, 78(5): 756–768.
[58] Talwar, S., A. Wiek, and J. Robinson. 2011. User engagement in sustainability research. Science and Public Policy, 38(5): 379–390.
[59] Walter, A. I., S. Helgenberger, A. Wiek, and R.W. Scholz. 2007. Measuring societal effects of transdisciplinary research projects: Design and application of an evaluation method. Evaluation and Program Planning, 30(4): 325–338.
[60] Wiek, A., S. Talwar, M. O'Shea, and J. Robinson. 2014. Toward a methodological scheme for capturing societal effects of participatory sustainability research. Research Evaluation, 23(2):1 17–132.
[61] Sheppard, S., D.M. Iype, S. Cote, and J. Salter. 2015. Special Report – A Synthesis of PICS-Funded Social Mobilization Research. Collaborative for Advanced Landscape Planning, University of British Columbia.
[62] Westerhoff, L., S.R.J. Sheppard, D.M. Iype, S. Cote, and J. Salter. 2018. Social mobilization on climate change and energy: An evaluation of research projects in British Columbia, Canada. Energy Research and Social Sciences, 46: 368–380.
[63] Chilvers, J., and M. Kearnes. 2016. Remaking participation: toward reflexive engagement. Pages 261–288 in J. Chilvers & M. Kearnes (Eds.), Remaking Participation: Science, Environment and Emergent Publics. Routledge. P. 40.
[64] Luederitz, C., N. Schäpke, A. Wiek, D.J. Lang, M. Bergmann, J.J. Bos, S. Burch, … F.R. Westley. 2016. Learning through evaluation – A tentative evaluative scheme for sustainability transition experiments. Journal of Cleaner Production, 169: 61–76.
[65] Schäpke, N., I. Omann, J.M. Wittmayer, F. van Steenbergen, and M. Mock. 2017. Linking Transitions to Sustainability: A Study of the Societal Effects of Transition Management. Sustainability, 9: 737.
[66] Coryn, C. L., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. American journal of Evaluation, 32(2), 199-226.
[67] CliNKS. 2014. Developing a theory of change (revised March 2014). https://www.clinks.org/support-evaluation-and-effectiveness/demonstrating-outcomes
[68] Coryn, C. L., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. American journal of Evaluation, 32(2), 199-226.
[69] Coryn, C. L., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. American journal of Evaluation, 32(2), 199-226.
[70] Global Affairs Canada - Results-Based Management for International Assistance Programming. 2016. Results-Based Management for International Assistance Programming at Global Affairs Canada: A How-to Guide. Second Edition.

[71] Coryn, C. L., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. American journal of Evaluation, 32(2), 199-226.
[72] https://www.nrcan.gc.ca/evaluation/reports/2015/18140#a3
[73] https://www.nrcan.gc.ca/evaluation/reports/2015/17155
[74] CliNKS. 2014. Developing a theory of change (revised March 2014). https://www.clinks.org/support-evaluation-and-effectiveness/demonstrating-outcomes
[75] Patton, M.Q., K. McKegg, and N. Wehipeihana (eds). 2016. Developmental Evaluation Exemplars. New York: The Guilford Press.
[76] Wong, S., G. Gaudet, and L.-P. Proulx. 2017. Capturing Wind with Thermal Energy Storage—Summerside's Smart Grid Approach. IEEE Power and Energy Technology Systems Journal, 4(4): 115–124.
[77] Wong, S., G. Gaudet, and L.-P. Proulx. 2017. Capturing Wind with Thermal Energy Storage—Summerside's Smart Grid Approach. IEEE Power and Energy Technology Systems Journal, 4(4): 115–124.
[78] CliNKS. 2014. Developing a theory of change (revised March 2014). https://www.clinks.org/support-evaluation-and-effectiveness/demonstrating-outcomes
[79] CliNKS. 2014. Developing a theory of change (revised March 2014). https://www.clinks.org/support-evaluation-and-effectiveness/demonstrating-outcomes
[80] https://www.communauto.com/index_en.html
[81] Celsor, C., and A. Millard-Ball. 2007. Where Does Carsharing Work? Using Geographic Information Systems to Assess Market Potential. In Transportation Research Record: Journal of the Transportation Research Board, No. 1992, Transportation Research Board of the National Academies, Washington, D.C., pp. 61–69.
[82] Braham, A. 2011. Contribution aux outils d'évaluation des impacts environnementaux de l'autopartage. Master's Thesis, École Polytechnique de Montréal.
[83] Sioui, L., C. Morency, C., and M. Trépanier. 2013. How carsharing affects the travel behavior of households: a case study of Montréal, Canada. International Journal of Sustainable Transportation, 7(1): 52–69.
[84] CliNKS. 2014. Developing a theory of change (revised March 2014). P. 26. https://www.clinks.org/support-evaluation-and-effectiveness/demonstrating-outcomes
[85] CliNKS. 2014. Developing a theory of change (revised March 2014). P. 26. https://www.clinks.org/support-evaluation-and-effectiveness/demonstrating-outcomes
[86] CliNKS. 2014. Developing a theory of change (revised March 2014). P. 26. https://www.clinks.org/support-evaluation-and-effectiveness/demonstrating-outcomes

Appendices

Appendix 1. Working Session on a Framework to Evaluate Low-Carbon Energy Experiments, June 1, 2018, Summary

CONTEXT. In Fall 2016, and in the context of their Generation Energy initiative, Natural Resources Canada (NRCan) commissioned Sustainable Canada Dialogues (SCD) to examine how Canada could transition its energy systems to low-carbon configurations while remaining globally competitive. The scholars' resulting report, *Re-Energizing Canada: Pathways to a Low-Carbon Future*, offered specific recommendations on how to advance the low-carbon energy transition. One of the recommendations for immediate action was to support low-carbon energy transition experiments in innovative social practices and technologies covering the spectrum of diversity found in Canada.

Indeed, the low-carbon energy transition entails not only the diffusion of new and existing technologies (e.g., the adoption of electric vehicles and new renewables) but also changes to the policy and regulatory frameworks (e.g., regulatory standards and carbon pricing regimes) to facilitate innovation and speed-up the adoption of novel social practices and norms (e.g., new perceptions of, and behaviors, around mobility and electricity use) (Geels & Schot 2007, Smith et al. 2005, Senbel et al. 2014). In the context of necessary transformational changes, widespread and diverse experimentation is essential, because we do not know exactly what the future will look like (Nilsson et al. 2011). The characterization of the costs and benefits of different approaches, their distribution over time and space, and their relative desirability and social acceptability cannot be fully known in advance. Learning, therefore, serves as a core component of processes of system innovation, particularly the iterative relationship between 'doing' and 'learning' as transition experiments begin to carry forward transformative change on the ground (Loorbach & Rotmans 2010, Gross 2010).

Rosenbloom & Meadowcroft (submitted) described **transition experiments** "*as projects that: (a) are explicitly linked to a larger vision of the long-term transformation of our energy systems to low-carbon; and (b) are consciously understood as trials to test the viability of innovations and to explore possibilities for change.*" Transition experiments can thus be understood as deliberate interventions that explicitly embody a novel configuration of social and technical elements that could lead to substantial low-carbon change in energy systems. Rather than simply demonstrating technologies or technological applications,

they act as small-scale trials of new models for meeting societal functions, shedding light on novel ways to produce and distribute energy, transport people and goods, or develop built environments.

On June 1, SCD therefore invited a group of experts with relevant experience from the private, NGO and public sectors (see Participants Section) to Ottawa to a one-day working session to follow up on the idea of developing an energy transition program as part of Canada's efforts to transition to low-carbon energy. The scholars held the working session to begin co-developing a holistic approach for evaluating socio-technical energy transition experiments and show how lessons learned can inform and accelerate Canada's low-carbon energy transition.

WORKING SESSION. The working session began with a brief presentation of the context of the workshop, in particular SCD's desire to accelerate the low-carbon energy transition, and a brief presentation by an SCD scholar of the key characteristics of transition experiments. Each one of the 24 participants then described an example low-carbon energy transition experiment with initial ideas about its potential evaluation (Output 1). Before the participants moved to breakout groups, Steve Williams, President of Constructive Public Engagement and Ph.D. candidate at UBC focussing on the evaluation of transition experiments, presented key elements of the evaluation framework for transition experimentation on which he is currently working. Each participant then recorded what they consider key characteristics of a robust evaluation framework (Output 2). They then joined one of three groups with representation from each sector to discuss the desired elements of an evaluation framework (Output 3). In a second breakout session, the groups attempted to develop an evaluation framework for one of three case studies of the low-carbon energy transition: (1) the City of Toronto's renewable natural gas projects, (2) the Sustainable Neighbourhood Retrofit Action Plan in West Bolton, Ontario and (3) Communauto, a Montreal-based car sharing service (Output 4). In a general discussion at the end, participants exchanged on key issues that emerged during the day and that could inform the development of a low-carbon energy transition experimentation program in Canada (Output 5).

TAKE-HOMES FROM DISCUSSION

Output 1: Examples of low-carbon energy transition "experiments"

Participants illustrated the breadth of low-carbon energy transition experiments using a range of example at various scales:

- individual level, e.g., a web tool to encourage individuals to adopt an emissions pledge in line with international climate goals and track their carbon emissions, which then provides suggestions tailored to their own emission profiles
- household energy efficiency level, e.g., sending homeowners infrared pictures of their house and testing whether it influences the home retrofit rate
- community level, e.g., training diesel-dependent communities in energy retrofitting, thereby creating jobs
- urban examples, e.g., building neighbourhood identity and engagement with the co-creation of a Sustainable Neighbourhood Retrofit Action Plan program across the Greater Toronto Area; using autonomous vehicles to improve neighbourhood liveability in Vancouver Surrey
- industry level, e.g., greenhouse growers using surplus heat from neighbouring industry activities; turning renewable-natural-gas-powered garbage trucks into mobile data units
- utility level, e.g., using real-time time-of-day pricing for electricity

- commercial level, e.g., the "transition" cereal of Kashi allowing farmers to sell "organic products" while awaiting certification
- fully-controlled, randomized and replicable design, e.g., the Nest thermostat in the UK
- country level, e.g., Canada's carbon pricing

Output 2: Key characteristics of a robust evaluation framework

Prior to further discussion, participants individually wrote down what they consider key characteristics of a robust evaluation framework. All responses were transcribed verbatim in Microsoft Word, and *a posteriori* codes were generated from the responses using open source TAMSAnalyzer software (version 4.48b2ahYos). This coding analysis was conducted on participant responses to identify the most common responses and also assess similarities and differences in responses across sectors (civil society and municipal government, university, and federal government) (see Table 1).

There was broad, cross-sector agreement around the need for an evaluation framework that focuses on transformative social and environmental change that enables a sustainable transition towards low-carbon energy, and the importance of scaling-up learning and motivating behavioural change. Governance was mentioned as a key characteristic to be evaluated, but details were not prescribed. Participants across sectors asserted that greenhouse gas emission reductions must be a specific desired outcome, and the framework must be created via participatory collaboration with its users and project proponents.

Civil society and municipal and federal government were clear on the need to evaluate and feedback learning over the full course of the transition experiment. There was agreement among the scholars and the federal government on the need to assess the full scope of desired outcomes, including built capacity and co-benefits, across a range of transition experiments, while assessing the context-dependence (and therefore transferability) of results. More specific characteristics of the framework, such as considerations of justice and financial returns, as well as the attitude towards failure and ability to adapt to changing external contexts were mentioned across, but not by all sectors.

Finally, participants from the federal government had a special interest in program design. There was repeated emphasis among federal participants for a clear framework developed at the onset of the transition experiment, that allows comparison to a counterfactual (such as a baseline or control group), using multiple, pre-defined data sources, and the need to consider how to navigate political circumstances.

Output 3: Discussion of the desired elements of an evaluation framework

Faced with the diversity of possible energy transition experiments, the participants agreed that a range of social, technical and economic transition experiments across different scales should be evaluated, from the community, city, industry and country level. They felt that transition experiments should fall under a portfolio led by various partnerships of stakeholders, including Indigenous, municipal, provincial and federal governments and non-governmental organisations. An effective framework, it was argued, will thus require having conversations with different sectors to understand their differing needs and circumstances as well as the limitations imposed by the "real world".

At the same time, participants agreed that not every project or innovation counts as a transition experiment. Rather, transition experiments must fit into a broader vision of the transition to low-carbon energy and be evaluated on how to best achieve the desired outcomes. There was also agreement that the project should be defined as a transition experiment from the beginning.

Table 1. Key characteristics of a robust evaluation framework as written down by participants. X = mentioned by 1 representative. **XX = mentioned by more than 1 representative. Shading represents the prevalence of the response across sectors.**

Key Characteristics of an Evaluation Framework	**Civil Society / Municipal Gov**	**University**	**Federal Govt**
Considers social and environmental impacts (sustainability)	XX	XX	XX
Addresses scaling-up	XX	X	XX
Enables learning	XX	X	XX
Evaluates transformative change/Enables the transition	X	XX	X
Considers/Evaluates governance	XX	X	X
Results in/Enables behavioural change	X	X	X
Has GHG reductions as a goal	X	X	X
Is co-created with users/participants	X	X	X
Includes feedback/evaluation throughout the course of the experiment	XX		XX
Considers equity/justice	X		XX
Considers financials/returns on investment	X		X
Is flexible	X		X
Is amenable to/positive towards failure	X	X	
Experiments on multiple scales			
Is adaptive to changing contexts	X	X	
Builds capacity		XX	X
Measures co-benefits		X	XX

Evaluates effects of context		XX	X
Focuses on outcomes		X	XX
Includes a range of experiments		X	XX
Enables collaboration/networking		X	X
Identifies goals from the onset		X	X
Establishes how outcomes will be measured		X	X
Informs policy change		X	X
Uses qualitative and quantitative data		X	X
Identifies risks		X	X
Compares to a counterfactual (e.g. baseline or control group)			XX
Identifies data sources			XX
Is developed early in project design			XX
Is clear			XX
Provides space for consideration of political/subjective realities/How to move experiment through politica process			XX
Is specific enough to compare experiments			X
Is values-based			X
Identifies experiment's assumptions			X
Identifies the target audience			X
Attributes causes to effects			X
Is objective/rigorous			X

Is comprehensive			X
Assesses counterfactuals/alternatives			X
Does not constrain the implementation/design of the initiative			
Requires a dedicated space for discourse on commodity industries	X		
Evaluates the role of supportive platforms	X		
Requires different tools for different scales	X		
Is transparent	X		
Data are owned by communities	X		
Addresses the degree of ambition of the experiment		X	
Has broad goals		X	
Considers Indigenous Reconciliation		X	
Considers public education		X	
Is restricted to certain kinds of experiments		X	

Participants asserted that this would require that the transition experiment's definition of success (i.e., its objectives) be defined from the onset, as should the hypotheses, methods, metrics for measuring whether the goals have been achieved, assumptions of the transition experiment, the target audience and the metrics for measuring whether the target audience is being reached.

Participants felt that the framework should:

- Be rigorous, with a baseline and/or control group to which to compare it. This includes comparing counterfactuals or alternative paths, i.e., identifying the opportunity costs of and risks associated with the transition experiment. Given incomplete information available, participants agreed that multiple data sources and a mix of qualitative and quantitative metrics would have to be used.

- Apply at the onset, as early as possible during transition experiment design, and over the entire course of the transition experiment. The framework should thus allow reflexive adaptation/course-correction/feedback as the external context around the experiment changes and as we learn from the transition experiment, evaluating the value created by the process as well as the outcomes.

- Be flexible and general enough to apply to various kinds of transition experiments with various definitions of success, but also specific enough to be relevant to particular transition experiments and compare them.

- Identify actionable recommendations and results that could be replicated, scaled up and generalised to enable learning between and from transition experiments, to raise public awareness, encourage engagement and behaviour change and build long-term capacity, and thereby accelerate the low-carbon energy transition. The framework should therefore aim to attribute effects to causes (i.e., attribution as compared to contribution to effects), teasing out what can be applied in other contexts versus what is context-dependent.

- Be a multi-stakeholder collaboration that is participatory/co-created with project proponents who might not have the capacity to undertake the evaluation themselves, and with eventual users of the framework. This approach helps build relationships, share ideas and enable dialogue between stakeholders, who begin the transition experiment with an understanding of shared values and visions.

- Be simple, practical and transparent, in order to be applied effectively.

- Evaluate the transformative potential of the transition experiments, i.e., whether it helps the transition towards low-carbon energy, as opposed to solely whether it results in incremental change. For example, the framework should assess whether the transition experiment leads to a narrative or systemic change. This requires a focus on outcomes, and not just outputs (products). The framework must therefore identify what are the values and outcomes that are desired. Participants felt that sustainability/sustainable development should be a clear desired outcome. Specific outcomes mentioned were: (1) social—concerning justice, equity, public engagement and adoption, behaviour change, liveability, neighbourhood identity, community cohesion, Reconciliation and policy, (2) environmental—with a specific emphasis that greenhouse gas emissions should be reduced, as well as (3) economic—for example in terms of returns on investment, in acknowledgement of the financial reporting realities of stakeholders. Participants stated that the framework should also identify co-benefits and unintended outcomes, and evaluate the long-term impacts after the transition experiment has ended as well as the global impacts, to ensure problems (e.g., emissions) are not simply exported elsewhere.

Output 4: Developing an evaluation framework for three case studies

In groups with cross-sectoral representation, participants attempted to develop an evaluation framework for one of three case studies. In response to the questions "What is the single most difficult aspect of evaluating the project?" and "What is the key elements of an evaluation framework that cannot be compromised?", each group described different challenges.

Group 1: The City of Toronto's renewable natural gas projects

Case Study 1: The City of Toronto's Solid Waste Management Services Division is implementing a low-carbon energy project that will generate renewable natural gas (RNG) in support of goals set by the City's Long Term Waste Management Strategy and the TransformTO Climate Action Strategy. A bio-methane upgrading system will allow the City to take the raw biogas produced by processing Toronto's Green Bin organics and upgrade it into RNG. The upgraded gas will then be injected into the natural gas grid for use as renewable fuel. The project supports the City's move toward a circular economy by using a closed-loop approach where organics collection trucks can ultimately be powered by the waste product they collect.

The discussion touched on the need to assess unintended consequences and the importance of separating actual transition experiments from projects that borrow the language of experimentation while ultimately reinforcing the status quo. A critique of the project from a "transition experiment" perspective is that its reliance on an existing Enbridge pipeline for RNG distribution might empower Enbridge to say they have achieved greener NG in their grid. It is unclear *a priori* what a municipality like the City of Toronto can do with such criticism, since it is not in a position to lead the creation of, for example, a fully electrified fleet of heavy duty collection vehicles and long-haul trucks for waste management, given the lack of supporting infrastructure and the lack of market-tested technologies when, in contrast, it has existing infrastructure/technologies and both environmental and financial business cases supporting our transition to CNG/RNG. Solutions that were proposed included preventing a monopoly supply in the hands of Enbridge, and finding industrial partners that could improve injection technology, as a new

niche for emerging companies. This discussion highlighted the importance of considering the various actors upon which an energy transition program could be applied. It also shed light on the real constraints that shape municipalities' work, when developing criteria to avoid excluding such projects from future consideration. Importantly it emphasized that lessons that can be learned from unintended consequences and that a transition context would then allow to address those.

Group 1 identified establishing counterfactuals as both the **most difficult and also the most key element that cannot be comprised** of an evaluation framework. Specifically, the challenge is in identifying an alternative path or transition experiment, and then identifying which path or transition experiment leads you to the desired outcomes in the best way. Group members stated that making this choice requires asking from the onset what puts the transition experiment onto the needed trajectory, i.e., how does the experiment fit into the broader vision of the low-carbon energy transition?

Group 2: The Sustainable Neighbourhood Retrofit Action Plan in West Bolton, Ontario

Case Study 2: The Sustainable Neighbourhood Retrofit Action Plan (SNAP) is a locally tailored home retrofit program reaching both homeowners and renters with building upgrades and community enhancement initiatives. The proposed West Bolton (Caledon) SNAP model is focused on downscaling the Provincial and municipal climate change plan down to the neighbourhood scale, and working to empower residents to take advantage of incentives to implement residential energy efficiency retrofits by using the strengths inherent in face-to-face interactions with a local delivery agent. Working together with residents, the municipality, and utility partners, we will co-create a home energy retrofit program

using tactics that are both "high-tech" (social media, online surveys, and other interactive tools) and "high-touch" (smaller scale in-person engagement using focus groups, workshops and door-to-door conversations). Through this process the West Bolton SNAP will contribute to collective identity creation in the neighbourhood, and the Town of Caledon, around issues of climate change and sustainability, and will support individual residents with taking action.

The bulk of Group 2's discussion was around the accessibility of data needed to perform an evaluation. Indeed, Group 2 felt that the lack of accessible data is **the most challenging aspect of evaluating transition experimentation**. Data on greenhouse gases are not readily available, nor are private data, such as those from private utilities; moreover, it is difficult to match the geographic scale of the experiment to the geographic scale of available data. Group members suggested that participants could be surveyed before and after the transition experiment or compared to a similar community to identify a more rigorous effect, however, survey data (i.e., opinions) may not be reliable. Moreover, a counterfactual neighbourhood could be identified as a baseline indicator, but the same data issues could exist in this other neighbourhood. Concern was also raised that the time and money invested in evaluating the project would reduce availability for project implementation. The group members also emphasised that any evaluation would have to address what each funder would want out of an evaluation. **The key elements of evaluation that cannot be comprised** for Group 2 were the creation of a proxy measure of outcomes for all relevant outcomes—how to define measurable metrics for desired outcomes—and having a clear definition of success.

Group 3: Communauto, a Montreal-based car sharing service

Case Study 3: Communauto is a Montreal-based car sharing company that operates in four cities in Québec, in Halifax, Kingston and Ottawa, and in Paris, with a fleet of over 1100 vehicles (as of 2011) and 40,000 members (as of 2015). The company was established in 1994 with the business model to provide a convenient, cost-effective alternative to car ownership that simultaneously reduces the environmental impact of private travel. Recognizing people's reticence to get rid of their personal cars, Communauto markets car sharing as complementary to city residents' habitual modes of transport. The intention is to reduce the environmental impact of transport by discouraging personal car ownership, but also encouraging active modes of transport. Communauto actively positioned itself as a key component in a city's transportation ecosystem, particularly in Montreal, through partnerships with other transportation organizations such as Montreal's public transit authority, la Société de Transport de Montréal (STM), Montreal's bicycle sharing program, BIXI, VIA Rail, and car sharing services in neighbouring provinces.

Group 3 echoed the challenge of identifying counterfactuals, and the practical difficulties of finding a similar case study to which to compare the transition experiment. For example, Communauto was founded in Québec City as a non-profit, where it failed, and then it was brought to Montreal as a for-profit, where it succeeded, but it is not clear whether this is the cause of success that should be replicated. In addition, group members asked: Is Communauto itself a car-sharing company, or a multimodal transportation system, given its partnerships with other transportation organizations? Defining the experiment would enable identifying the proper counterfactual. It followed that, for Group 3, **the most challenging aspect of evaluation** is the identification of what factors lead to success: What are the policy and supporting components needed to replicate successes?

Output 5: Points for further discussion

Replication and controlled experiments vs. transition experiments

Participants discussed how to translate the desire, expressed in particular by participants from the federal government, for the "gold standard" of experimentation, as in randomised control trials, to the complex real-world systems of transition experiments where such controlled settings are not always applicable. It was stressed that experimentation implies an effort to assess the impact of an initiative by means of a structured comparison. Traditional experimental thinking understands that comparison is to a so-called baseline or some other control group that functions as a counterfactual. In the language of experimentation, the counterfactual is used to tell us what would have happened otherwise. It was suggested that this definition is broad enough such that, even if we are using only qualitative evidence and case studies as methods, we can still approach the evaluation with this framework in mind. It was also pointed out that numerical ecology provides rigorous statistics that can be applied in in social-ecological systems where observational studies are conducted. Participants reiterated the need to mix quantitative and qualitative data and the need for methodological pluralism in transition experimentation.

Scaling up

The difficulty of transferring and scaling up lessons learned to other projects was discussed. This included how to make sure the needs of economic actors and industry are heard by the federal government to enable scale-up and policy outcomes. Participants identified that different project proponents will look for different outcomes from experimentation, and there is therefore a need to be inclusive in the design of the framework. Participants presented the model of labs or cohorts to bring stakeholders together and create a safe space for sharing ideas, capacity and data, fostering synergies between experiments. It was also noted that even if we know what led to the success of an initiative, such as Communauto, that is not necessarily what we need to replicate to create additional successes/transitions. For example, entrepreneurs are now replicating the successful model of Airbnb, or at least attempting to replicate it in other contexts, and it is much easier because they simply have to say: "I am establishing Airbnb for warehouses, or Airbnb for tents, or Airbnb for X". A successful experiment actually makes replication easier.

Experiment or pilot or initiative or innovation?

Participants discussed the appropriate terminology for experimentation, mentioning the need to frame and explain it properly, given historical misuses of "experimentation". It was mentioned that "experiment" does not resonate at the municipal level because it is not a term that municipalities tend to use ("pilot" being the obvious alternative). Although municipalities run trials and tests, a participant suggested that it was unlikely to find a municipal project that would like to be identified as "experiment", in large part because governments are expected to make responsible rather than risky use of public funds and resources. Furthermore, the social/political dynamics of the word "experiment" are an essential consideration for any organization—and in particular municipal governments—whose work has a real and direct impact on people. Bureaucrats have a terrible legacy of running failed experiments on vulnerable communities—many of which still live with the consequences—and we should be respectful of that history.

Learning from failures

The discussion on risks and failures expands on the previous paragraph. Participants discussed the need to approach experimentation with a mindset oriented towards learning, arguing that it follows that much had to be learned from failure or from unintended consequences of an initiative, which should not be confused with the methodological failure of designing a flawed evaluation. It is important to have rigorous evaluations for learning to occur, even when the initiative that is evaluated does not live up to

its initial promise. The need to consider the disparate effects that failure or unintended consequences can have on different groups, depending on their social role and the costs and benefits to them was also mentioned since real individuals will be directly involved/affected by a program of energy transition. A participant highlighted the need to consider what types of projects/opportunities the framework might exclude if a rigid definition of experimentation is adopted calling for additional/supportive conditions to create a permissive environment for risk-averse actors like municipalities to do the kind of transformational work discussed during the working session.

For additional information: catherine.potvin@mcgill.ca

PARTICIPANTS

Beauregard Tellier, Frédéric	Exc. Director, Pan-Canadian Framework, Environment and Clima Change Canada
Chesterley, Nicholas	Manager, Innovation Unit, Labour Program, Employment and Soci Development Canada
Clark, Ian	Acting Chief, Economic Research, Science Integration and Outreac Strategic Policy and Results Sector, Natural Resources Canada
Davis, Meaghan	Senior Project Manager, Strategic Projects, City of Toronto
Etzion, Dror	McGill U., Sustainable Canada Dialogues
Hughes, David	President & CEO, The Natural Step Canada
Kalin, Michael	Experiments and Analytics, Strategic Policy and Results Sector, Natur Resources Canada
Kenney, John	Office of Energy Efficiency, Natural Resources Canada
Khan, Haris	Impact & Innovation Unit, Privy Council Office
Lamarche, Mary Kay	Evaluation Manager, Natural Resources Canada
Leonard, Niamh	Program Officer, McConnell Foundation
Leyburne, Drew	Director, Strategic Policy and Results Sector, Natural Resources Canac
Lysack, Mishka	U. Calgary, Sustainable Canada Dialogues
Mathew Iype, Deepti (remote)	U. British Columbia, Sustainable Canada Dialogues
Matthews, Damon	Concordia U., Sustainable Canada Dialogues
McCurdy, Patrick	U. Ottawa, Sustainable Canada Dialogues
McKenzie, Janetta	U. Waterloo, Sustainable Canada Dialogues
McVey, Ian	Project Manager, Ontario Climate Consortium

Meadowcroft, James	Carleton U., Sustainable Canada Dialogues
Monafu, Dan	A/Senior Advisor, Priorities and Planning, Strategic Polic Experimentation, Treasury Board Secretariat
Mousseau, Normand	U. de Montréal, Sustainable Canada Dialogues
Poirier, Eric	Director, Smart Cities Challenge, Infrastructure Canada (present in tl morning only)
Potvin, Catherine	McGill U., Sustainable Canada Dialogues
Rahbar, Shahrzad	President, Industrial Gas Users Association
Sanchez Valero, Caroline	Vice President, Réseau Environnement
Stoddart, Mark (remote)	Memorial U., Sustainable Canada Dialogues
Williams, Steve	President, Constructive Public Engagement; University of Britis Columbia

TO BE INVOLVED IN SUBSEQUENT DISCUSSIONS:

Cabaj, Mark	President, From Here to There Consulting
Gélinas, Johanne	President, Transition énergétique Québec
Langelier, Jean-Sébastien	Senior Analyst, Infrastructure Canada
Perla, Keran	Director, Foresight and Design at the Department of Energy, CoLal Prov of AB
Ryan, Alex	VP, Systems Innovation, MaRS Discovery District
Sadik, Tonio	Director of Environment, Assembly of First Nations
Singer, Janice	Industrial Research Assistance Program, Natural Resources Canada
Taillefer, Alexandre	President, XPND Capital
Vallillee, Andre	Environment Program Director, Metcalf Foundation

Appendix 2. Discussion of Evaluation Framework for Experimentation: City of Summerside and Sustainable Canada Dialogues, Summerside, PEI, July 20–24, 2018

Table of Contents

Participants

City of Summerside:

Bob Ashley, CAO
Gordon MacFarlane, Deputy CAO
Greg Gaudet, Director of Municipal Services
Aaron MacDonald, Director of Technical Services
Gerald Giroux, Electrical Engineer
Mike Thususka, Director of Economic Development
Rob Philpott, Director of Financial Services
Bobby Dunn, Program Manager, Heat for Less Now

Sustainable Canada Dialogues:

Divya Sharma, McGill University
Matthew Hall, University of PEI

Meeting Agenda

July 20th: Bob Ashley and Divya Sharma
July 23rd morning: Gordon MacFarlane, Greg Gaudet, Aaron MacDonald, Gerald Giroux, Mike Thususka, Rob Philpott, Bob Ashley and Divya Sharma
July 23rd afternoon: Greg Gaudet, Gerald Giroux, Matthew Hall and Divya Sharma
July 24th: Bobby Dunn and Divya Sharma

Sustainable Canada Dialogues Summary

In Fall 2016, and in the context of its Generation Energy initiative, Natural Resources Canada (NRCan) commissioned Sustainable Canada Dialogues (SCD) to examine how Canada could transition to a low-carbon energy economy while remaining globally competitive. The scholars' resulting report, *Re-Energizing Canada: Pathways to a Low-Carbon Future*, offered specific recommendations on how to advance the low-carbon energy transition. One recommendation for immediate action was to support "low-carbon energy transition experiments" in innovative social practices and technologies covering the spectrum of diversity found in Canada. A "transition experiment" can be understood as any project involving social and/or technological innovations that could lead to substantial low-carbon change in energy systems, where critical evaluation to extract lessons is explicit. In May 2018, NRCan commissioned SCD to produce a 20-page report on a framework to evaluate low-carbon energy experiments, due in mid-September 2018. On June 1, SCD thus invited a group of experts with relevant experience from the private, NGO and public sectors to Ottawa to a one-day working session to begin co-developing a holistic approach for evaluating socio-technical energy transition experiments and show how lessons learned can inform and accelerate Canada's low-carbon energy transition. To co-create the framework in a participatory process, the scholars will work with project proponents from four case studies, including the City of Summerside's smart grid project.

City of Summerside's Smart Grid Project Summary

In 2009, Summerside Electric, the municipally-owned electric utility of the City of Summerside, Prince Edward Island, installed a 12 MW wind farm and contracted an additional 9 MW of capacity from

another wind farm. As this constitutes a significant amount of wind energy relative to Summerside's electricity demand, which peaks at 28 MW, Summerside Electric began a smart grid initiative to help make demand better match wind availability. These initiatives resulted in the notable achievement of wind energy providing 50% of the city's annual electricity consumption.[1]

Summerside's smart grid uses fibre optic and wireless communication to exchange real-time data between the utility and participating customers. This facilitates smart heating and hot water appliances which respond to grid signals. Summerside Electric's Heat for Less Now (HfLN) program sees residents and businesses replace existing heaters (typically oil furnaces) with electric thermal storage devices that combine resistance heaters with heat-storing ceramic bricks. The program also includes electric hot water heaters that can vary in temperature. Together, these appliances provide a means of adjusting electricity consumption by storing thermal energy, a capacity which is controlled by the utility in real time. Participation in the program is incentivized through discounts on the electricity consumed by connected appliances, provincial rebates and options to lease.[1] Participants receive feedback on appliance electricity consumption, costs and GHG emissions from an online customer portal. By shifting demand to time periods when more energy is available, the HfLN program provides financial benefits for both the utility and consumers.

Discussion Summary

The goals and considerations of City of Summerside projects like MyPowerNet and Heat for Less Now are largely economic; greenhouse gas emission reductions and related environmental results are considered valuable co-benefits, but not the primary objective of any program. Evaluation is considered useful for market validation of innovations, but also internally as a decision-making tool to provide feedback on the process and outcomes and whether a project is financially worth the investment. It is also deemed useful as a way to enable consumers to share project successes with others and raise project awareness. Some senior city management stated that standardised, independent, third-party evaluation is preferable, as it ensures standard performance metrics for the industry are used and that the evaluation is reliable and unbiased. They mentioned that evaluation would have to be built in to the project plan from the onset, to incorporate evaluation costs into project budgets. Both quantitative and qualitative data would have to be collected to measure indicators like energy usage and savings, uptake rates, reduced peak demand, customer behaviour and more. Successes of the program were attributed to the fact that the City has the technical expertise to make informed decisions, and full control over the utility and City assets. There was disagreement among city officials in terms of the replicability of Summerside's model in municipalities that do not own their own utility. Some felt that Summerside's ownership of its utility would present a confounding factor in a comparative evaluation; others claimed that the Summerside model could be replicated since it is fundamentally about collaboration between a utility and a city. At the moment, evaluation of City projects is *ad hoc*, due to lack of time, skilled staff and importance placed on evaluation compared to more immediate concerns.

Discussion Minutes

Key Indicators/Metrics of Success

Cost savings of participating households

[1] Wong, S., G. Gaudet, and L.-P. Proulx. 2017. Capturing Wind with Thermal Energy Storage—Summerside's Smart Grid Approach. IEEE Power and Energy Technology Systems Journal, 4(4): 115–124.

GHG emission reductions of participating households

Uptake by residents of new technologies or behaviours

Rate of uptake by residents of new technologies or behaviours, i.e., uptake over time, not just volume

Weekly engagement/interaction of the City with interested households

Load factor compared to furnace loads

Energy production at site-specific locations

Utility income (energy sales)

Reduced wind energy surplus—use of excess wind energy internally, vs. spilling into the grid or selling

Reduced peak energy demand compared to previous years

A less than five-year payback for technology, so that the decision for the household to participate is palatable

Effect of program on other rate payers or on budgetary constraints, i.e., introducing a smart grid in Summerside cannot be at the cost of something else—it has to be self-sustaining financially (it has to pay for itself).

Reliability of the system, i.e., the project should not impact the City's current, day-to-day operations. A new system should not stop us from flicking a switch to turn on a light or make that more complicated.

An indicator of the attractiveness of Summerside to entrepreneurs approaching its Living Lab

Uptake of the benefits of provincial demand-side management programs by citizens and businesses

Qualitative and quantitative data on customer satisfaction with the City's provision of data to them, i.e., Is it helpful for you to know how many GHGs you are creating compared to your neighbour? Does it make you feel good that you have made the choices you have? Are you being rewarded for the choices you have made? Cognitive behavioural economics and psychology are intriguing as the new emergent field of policy development. What motivates people to change their behaviours? It would be interesting to see if friendly competition or peer pressure would influence individual consumer choices, asking, for example: Did you know that ¾ of the people on your street are producing 1000 tonnes less GHG emissions than you are? Feedback tools are important to show consumers the effects of their behaviour; the City needs to understand how this feedback will make the consumer feel good or important.

Evaluation Process

What are the objectives? The reasons to evaluate/value of evaluating?

As a proof of concept. Market validation is critical for any business trying to grow a footprint in North America. As one introduces innovation, and there is resistance to change, having market validation is critical.

As a decision-making and a forecasting tool. To know whether to pursue, abandon or modify the project, and whether to continue investing in the project or not.

To inform the City's own procurement habits. Evaluation will help us assess the effectiveness or efficiency of procured equipment, software, and other supplies related to any project.

For communications and outreach and to influence consumer behaviour with documented support, including political messaging and policy development on the administrative or bureaucratic side. Political messaging is equally as important in mobilising the population towards making better choices. To increase consumer awareness and promote the project as participants share successes with non-participants.

As an integral part of grant funding applications.

For whom is the evaluation intended?

Both internally and externally, depending on the project and who the required partners are: the City Council, federal/provincial agencies, city officials.

City administrators, senior bureaucrats and staffers in economic development, our electric utility and recreation/leisure.

The entrepreneurs who approach the City's Living Lab, in terms of raising capital and sales.

Who is carrying out the evaluation?

From the perspective of entrepreneurs who approach the City's Living Lab, third-party, unbiased evaluation is critical. If the City was seen as both the proponent and support of the experiment, it did not work out. Additionally, salesmen come to offer the City innovations, but they need third-party performance metrics to prove what they do is true. The City has enough internal expertise to say if a project should move forward, but sometimes city officials have rose-coloured glasses. If the intent of evaluation is to sell more product, the economy or industry expects third-party evaluation or, at least, standards. A third-party approach is imperative because it removes bias; most people would not move from a path once they have committed to it.

The City does not have an evaluation department. Evaluation would thus be a distributed, delegated function of perhaps any department that can use the evaluation.

Who provides the funding for the evaluation?

The City would fund the evaluation, if it is useful to them. However, if it is contributing to a national agenda, and if it is expensive, the City would look for federal assistance. The City's primary function is to provide services, not to do evaluative research. Their evaluations are mainly *ad hoc* and anecdotal, which is their weakness.

Evaluation would have to be put into the project budget at the beginning as an expectation of cost.

What lessons learned from evaluation can be shared with others externally? Are there lessons learned that could be shared to accelerate uptake of similar programs?

Realistically, the City will share lessons if triggered by a request; it would not be self-directed. People hear about what the City is doing, they make the inquiry, and the City is always forthcoming.

How are the results from the evaluation shared?

Electronically, either via internet or electronic reports. If they were fit for public consumption, the City would make the results available on its website. They would be shared with people like SCD. As long as they are intelligible, and not just data—the results of evaluation need narrative explanations and visuals. The City has shared a lot with its peers—other municipalities, associations, engineering groups, professional affiliates. The evaluation would have to be digestible, intelligible and utilitarian at the simplest, most basic level. Unless it is readily intelligible by the average person, it is going to be an impotent exercise. The City would share evaluations if they were available with federal or provincial funding partners, if the City were making an application for any type of energy-related project.

How can the results from the evaluation be integrated in project implementation?

Internally, the City shares the evaluation to prove whether its assumptions were correct and suggest whether the investment was worth it. If the City is working with entrepreneurs who approach the Living Lab, the entrepreneurs would share the results with a future market.

Evaluation would influence the City's marketing and promotion of its programs. Evaluation could be threaded into and become integral to any communications dealing with the program. Evaluation could change the City's procurement habits in acquiring supplies and equipment. Evaluation could affect the City's human resource estimations of what the City needs in terms of expertise. Evaluation could be useful for future policy development. Evaluation could be useful as an arm of innovation, as a way to help inform paths ahead and pitfalls to avoid.

What methods would/do you use? What data are available?

What will be recorded must be set up and the methods incorporated at the start—it is too difficult afterwards.

There is an internationally accepted evaluation protocol for energy usage data, to determine and measure energy savings. The method requires one-year minimum of data for whatever is being evaluating, which is a fairly significant effort. However, 18 months to two years would be better, as energy profile uses, driving habits and market habits change season-to-season. Some people ask for one-second datasets—that is difficult to get.

Evaluation has to include a societal, economic and environmental review.

Data: financial records, procurement records/purchases—number of units sold, rented, in operation; number of complaints; number of service trips (quantitative). Customer, management and political feedback (qualitative). Comparative data with non-renewable energy, in terms of GHGs saved, actual kWh consumed of electricity. Seasonal shifts in consumption, peak demand/peak loads, technical failures.

These data are available from work orders, purchase orders, sales invoices, electronic measuring devices/meters. Communications, emails, phone calls, surveys would be required for customer feedback.

The evaluation would become a toolkit for communications and public relations. Therefore, theoretically, the *analysis* for evaluation would be done automatically, with as much of the data collected through electronic means as possible. The City has a rich data field available to customers on the customer portal, in terms of how much energy their appliance is using, how much energy is being supplied by their windfarm, a comparison of GHGs they are not creating by not using oil, etc.

What conditions are needed?

'Big data' is up and coming and a critical tool. We are all busy; it would be helpful to simply have a green light indicating that the system is running smoothly. With a proper data management system, there is no reason why it cannot alert the engineers.

Electronic measuring devices, many of which we have, but others that are available that we do not.

Communication channels that are easily accessible and sharable, within the City and with external sources.

Financial structures or analysis.

Transparency. Evaluation could be used as a communications tool, revealing the public and other stakeholders, the pluses and minuses of any "experiment". Transparency is the clarion call of the media and citizens vis a vis all governments. It would not be our intent hide evaluation documents. Evaluative content may well need to be edited to make it intelligible to lay people.

Adequately trained staff to both produce and interpret the analysis. This includes procedural training in systems, repeatable procedures and accountability measures, like any other management function. As far as an evaluation process, a group with specialities is required—the technical aspect requires engineering expertise, the societal requires communication expertise, the environmental requires sustainability expertise. It is critical to have a board of different backgrounds to notice things that others would not.

What have you learned as the program has evolved?

Ad hoc, infrequent evaluation is being done and is usually event-driven. For example, a complaint or error may drive an evaluation. This evaluation is in the form of sharing data or feedback, meetings to discuss the data or the feedback and possibly an assignment to look into alternatives or causes of whatever has triggered the evaluation, like failures or deficiencies. Sometimes successes are also evaluated in terms of uptake. The City tries to figure out what may be changed from one state to the next. They do program reviews on a regular basis, but most would have nothing to do with energy; they would be in municipal programs like swimming or arts.

Challenges to Evaluation

Managing the public's expected outcomes. The utility's expected outcomes might not align with the public's expected outcomes. The utility built a wind farm but did not reduce energy costs. Therefore, should the evaluation criteria include a public-relations, communications aspect to set those expectations? The transition to low-carbon is slow, not immediate; and people have an expectation of

immediate outcomes these days. It can be a challenge communicating the broader benefits of what the City does to Canadians.

Data. How do we actually know how many kilometres each of us drives? How do we know at what temperature each of us has their houses? The City is relying on someone else's data.

Data collection will always be an issue—it is a lot of work. Baselines are difficult to achieve. Every community is different in one way or another. The number of variables and the complexity introduce all kinds of confounding conclusions. Interpretation of data could be an obstacle, because it requires skills, expertise and experience to interpret an evaluation in the most useful way. Also, it would not be easy to translate evaluations into strategic support mechanisms, i.e., we have a strategy to do X, how will the evaluation accelerate/strengthen/ support/complement our strategy?

Skills required. The City does not have the level of sophistication necessary for thorough, skilled evaluation. Evaluation is also easily overlooked—the pressures of scheduling, getting things done, and the fact that evaluation comes at the end makes it easily overlooked when the City has concerns about delivery, shipment and completion. People are anxious to wrap things up and expedite things. Evaluation would be a slow-down in some cases, as important as it is.

The City can supply a lot of data but does not have the horsepower to do an in-depth technical analysis.

Costs. The City had to monitor birds and bats after establishing the wind farm, so they budgeted for it upfront. It might be difficult to know the cost of evaluation beforehand.

Constraints from budget cycles. If the project involves federal funding that must be spent by a certain date, it does not always fit in the timelines of data collection needed—so budget cycles are a challenge.

Timelines. Sometimes evaluation requires a certain time period, so it can be difficult to incorporate into project design.

The need for commonalities and standards. Sharing on a broad scale may be difficult if evaluations do not have some kind of commonality or common standards that people agree on and share. It could be a real challenge to have this type of evaluation infiltrate or penetrate other academic disciplines that take up interest in this.

People started talking about smart grids 10 years ago, and they still have not released smart grid standards yet. The problem is that there are constantly new innovations that need to be incorporated. Luckily, the City did not sit and wait.

Industry associations, automobile associations need to play a role within evaluation. A standardisation is needed.

Penalties for early risk-takers. If the funding comes along and others jump on board, they do not have to do that in-depth evaluation that the City did. In other words, the City is ahead but penalised for it sometimes. The government eventually incentivises behaviour change and, if the City has already changed, they do not get the incentive.

Other Considerations for Evaluation

Evaluating future opportunities. A smart grid is evaluated based on its capability—it is a future-looking evaluation. Not only does one need to know that the investments one has made are doing well, but also what opportunities are coming forward. On the smart grid, the City has not yet asked itself what they can do with this, what are the technologies coming down the line that they could use, like an evaluation of opportunities. The City has discussed smart chargers and smart homes, but what are they going to do with it? Where should the next effort be? The City started controlling home appliances, but then found that there are many other appliances that can be controlled—what else is there that the City can look at? The smart grid was built to address a specific problem in the organisation—they have too much wind, so they need to manage it. It could be that the City is readying Summerside for electric or innovative transportation, and this is one way to address it. One thing the City was missing, that was brought by students from the University of New Brunswick who approached the Living Lab, was to use student resources to ask: What is the opportunity for all the transportation in Summerside? What would it look like if the City electrified every bit of it, assuming the electricity is from renewable sources? What is the economic value proposition, the role of the utility? There is opportunity there. Then the data are taken to policymakers to show the resulting savings on gas and increased sales for the utility to support city infrastructure.

Regional context/culture. PEI is an island—this makes it quite different. It is also the smallest province by an order of magnitude in Canada. The province is very in tune with its rural roots, and aware of its own reputation as a jewel in the ocean—there is romance here for tourists. How would you connect that to an evaluation tool? If some facts about the island culture are ignored, the study would be coloured by that. It is hard to compare PEI to other provinces that are not islands.

Part of the low-carbon transition is about education and how you evaluate that.

There is a lack of standardisation for the low-carbon transition, but that is what experimentation is for. Canada is starting wind programs, but does it have a baseline? Should we talk about information we should be gathering, so that when we fund these initiatives we can evaluate them 5–10 years down the road? Having baseline data and evaluating is critical, but you almost need an evaluation of an evaluation—if you have 20 GHG per person and then you want to get down to 15, does that mean I walk to work or drive an electric vehicle? How do you translate it into real life behaviour?

The Importance of Economics

The City uses cost savings to motivate the transition—officials present to the city council how much money can be saved. The City will do it if the economics make sense; it is all about finances.

The City's projects have forged ahead without clear objectives. The City follows the money and the funding sources and does things for the economic benefit of the community. GHG emission reductions occur because of that. An indicator of success is that the City uses more of its own wind energy at home, but if tomorrow they could sell it for a better price elsewhere than they would. The green initiatives that the City has worked on have made economic sense, as opposed to having an altruistic motive.

It is mistaken to ask: What has inspired people to be green? What is the social psychology of Summerside? The City established the smart grid to be completely practical. Summerside launched the wind farm because it needed a new ice rink and could not afford it, and it happened to have an electric

utility and program at the time that would fund the ice rink with the profits. Summerside is not all environmentally friendly; it is no different from any other community in Canada. It is very practical.

The consumer is fairly uneducated, and not looking at long-term planning, therefore the City offers rental and lease of smart appliances, so they do not have to pay out-of-pocket right away. There is a niche market that cares about its energy footprint when building a home, but it is small. Most people do not care about their GHG emissions; they just want their lights to work.

The City bought electric vehicles to influence the market to support the industry. It took a leadership role as the government, but that role has not filtered down to the community level. The City did it because it made economic sense for us, realising that this is more sustainable than always relying on government subsidies to support entry. The City's evaluation criterion is different from the consumer's—a connection has to be made to bring it to a mutual outcome of what both desire.

Consequences of Owning the Utility

Sharing objectives between departments. Provinces might not have a crown provincial utility and have not integrated well between government and utility objectives. Summerside has been able to take the energy industry and work with the municipality on other objectives to make sure they work better. That is part of the City's success, which is partly because they do not have a different shareholder that makes other demands. It is tough to innovate Canada's legacy systems when the shareholder is asking for a certain return on investment every year.

The City's infrastructure and history of innovation. There is an attraction of external business interests to the City due to its leadership. These businesses come for the City's infrastructure and historical nature of innovation, whether that has been purposeful or accidental.

The City's technical expertise. Policymakers need good technical bases for understanding what they are working on. Without the engineering behind it to do cost modeling and create graphs, one will not even get to that phase. For example, what are the actual numbers, cost, payback, fleet management policy for electric vehicles? Interest in electric vehicles has been driven socially, but without the engineering behind it, a city cannot get into it. Summerside is ahead of the curve, because the City looks aggressively at the economic model to monitor when the technology becomes sustainable.

Disagreement between senior management around replicating the Summerside model.

If there is any confounding in evaluation, it is that the Summerside model cannot be replicated across Canada. The governance of the utility is quite unique; it does not fall under the regular regulatory framework of most utilities. Everything that comes out of evaluation opinions will have to be qualified by the fact that Summerside has control over its own electric grid. It can make choices that 99% of communities cannot make.

The Summerside model can be replicated—it is about collaboration. The only difference is Summerside's collaboration is internal. Others can also work together, but it is about breaking down barriers and collaborating. Hopefully, other municipalities would be able to see that Summerside and the utility work together—they would talk to Summerside and have their utility talk to Summerside's utility to achieve an outcome.

Summerside is experimenting to try to grow the economy, and some of those components can be replicated, like Stash, Heat for Less Now and other programs.

Elements within the programs can be replicated, but the ecosystem in Summerside cannot.

Other municipalities do not have the capacity to evaluate.

But it might not be another municipality making the decision. In each experiment the City conducts, evaluation is based on the economic potential for investment in the community and how it makes the City more efficient to help other parts of Canada be more efficient too. A company like Samsung came to Summerside; they are trying to replicate it. Charlottetown and Summerside have different ecosystems, so they are successful in different ways—theirs is based on the vision of the provincial government, Summerside's is based on the City's, but they could achieve similar outcomes to each other.

It has taken so long for wind farms to be accepted—the utility wants to maintain reliability. Because of that they are wary of new technology. Summerside provides that gap, because it owns its own utility and city officials are pretty technical. All the things they have already adopted would have taken longer to be accepted in other cities. Summerside can be given as an example to municipal governments to try to have the utilities talk to the City on that side of things.

But think of the resistance to change meter!

Comparison with Charlottetown as a Baseline

Charlottetown as an appropriate baseline for Summerside. Charlottetown is a good comparison, because they have a lack of collaboration between different assets. Therefore, a comparison between Summerside and Charlottetown would be a comparison of a case where they do not work together versus where they do, within the same province. The mindset is relatively consistent across the island; consumers are similar.

Other municipalities that own their utility as an appropriate baseline for Summerside. Generic comparatives would be more appropriate, rather than specifically comparing one community to another—although there is the convenience of Charlottetown and Summerside being close but different. Summerside could be compared to a national average, to norms in the Maritimes—this is a much larger basis of comparison than strictly one city. It would be interesting to compare Summerside to Berwick, NS, or Anagonish, NS, because both have their own electric utilities, and therefore the same weapons. Municipalities that own and operate their own electric utility is a much more apples-to-apples comparison. Only in a municipality with its own utility would those matters become content for public debate and engagement and budgetary decision-making. Municipalities without electric utilities cannot create visions, strategic plans, policies, set budgets, prices or direct the activities of an entity they do not own.

A community owned electric utility presumes public engagement and decision making. This is not the case with privately owned electric utilities.

The demographics of the cities are different—higher density means lower energy consumption. This makes a comparison difficult. Summerside and Charlottetown also have different industries, for example Charlottetown has a coal burning plant.

Metrics to compare Summerside and Charlottetown

Space heating cost per capita

Savings from switching from heating oil to electric heating per capita

GHG emissions per capita

Penetration of renewable energy

Electric vehicle or electric lawnmower uptake

Electric heat vs. oil heat

General electricity consumption per household

Active transportation

Uptake of provincial energy efficiency programs—Charlottetown residents might insulate their homes more or buy more window caulking. What would that say about whether Summerside's program was successful or not?

Transit ridership

To attribute the differences between Summerside and Charlottetown to the smart grid, we would have to look at per capita data prior to and then during the smart grid period.

If we find a pattern of difference in those metrics between Summerside and Charlottetown, what would you attribute it to? What kinds of data would you look for?

Control of utility/assets. The City of Summerside controls the assets—Charlottetown has less control. Summerside owns its own utility and Charlottetown does not, and Summerside has introduced renewables, using electricity to offset other energies that are not as green.

Autonomy, technical leadership and access to decision-makers. Summerside is more integrated into its assets, whereas other municipalities are not. This means Charlottetown would have to have more collaboration, which leads to different agendas. Thus, the City's autonomy and technical leadership would explain the observed differences; Charlottetown does not have the technical leadership and staffing that Summerside does. For example, Summerside, being half the size of Charlottetown, has six engineers, while Charlottetown has three. The City's behaviour is changing due to its autonomy, flexibility and staff input to decision-makers.

Having the utility embeds expertise and experience and progressive political and administrative leadership that almost all other municipalities do not possess. City officials know how electrons flow, how to repair and control things, what the science and technology behind smart grids really is, how to install. The City is providing an essential service, so its level of electrical engineering is top flight.

Public awareness of programs that Summerside offers that Charlottetown does not. Citizens in Charlottetown do not even have the availability to make the choices that Summerside's citizens can.

Public awareness of provincial or federal climate change and energy efficiency programs.

Demographic strata—age, gender, marital status, ethnicity, education, income.

Basic population size. Could evaluation data be indexed or correlated to community size?

A comparison of Summerside to a generic town with an electric utility would require delving a little deeper into the energy programs and services that those municipalities offer or are engaged with.

Existing Data in Summerside

Five-minute-interval data on the health of controlled appliances like room heaters or domestic water heaters in participating homes, i.e., current brick and outside temperatures and maximum and current energy storage

Usage (kWh) over time per smart meter, down to five-minute intervals for commercial, but per-hour data for residential meters

Current windfarm output (power generation)

Current power rates vs. oil, propane and residential electric

Demographics of participating households: number of people at home, age, occupation

Current and historic energy mix of participating and interested households (~1–2000 households): electricity, oil, propane and wood—costs and usage

GHG footprint and savings per participating household

How interested individuals heard about the project (word-of-mouth, newspaper, radio, mail, social media, etc.)

Thoughts on Alternatives Paths/Futures

How to exceed expected outcomes? Along with the objective and the baseline, there is a baseline of expected outcomes, for example, one expects a certain number of units of electricity from the wind farm, based on the engineering and manufacturing information that one has. The question then is how to look at innovations to exceed that baseline of expected outcomes? There are many ways to look at it, either as a scientific, engineered outcome or something totally off-the-wall and innovative.

Identifying constraints. It is absolutely important to look at alternative paths; hindsight is always 20/20. The biggest burden to the smart grid was communications infrastructure. That choice of communications infrastructure was made for a more holistic view to lower telecommunications costs for the community, rather than it being the most effective use for a smart grid. Summerside is remote and has two telecom offerings for home services, so competitiveness is very lacking. People pay a lot more than they would outside. So the City tried to introduce a third competitor to lower pricing—this is why they started with fibre communications. The City said it needs a smart meter for the system to work well—those are also expensive. They have slowly adapted; they have a longer-term program for smart meters.

The need for a framework to review alternatives. On the technological review of alternatives, one has to look at everything and come up with a tool framework for how to look at alternatives. Then people could take the document and understand which sectors to look at.

There are other GHG reduction programs—Summerside could be investing in something that has minimal impact. The City is not against redirecting its path.

Appendix 3. Discussion of Evaluation Framework for Experimentation: Toronto and Region Conservation Authority (TRCA) and Sustainable Canada Dialogues (SCD), Vaughan, ON, 12–4PM, July 11, 2018

Table of Contents

Participants

Ian McVey, Toronto and Region Conservation Authority
Sonya Meek, Toronto and Region Conservation Authority
Jose Manuel Torcal, Toronto and Region Conservation Authority
Shannon Logan, Toronto and Region Conservation Authority
Adriana Gomez, Toronto and Region Conservation Authority
Divya Sharma, Sustainable Canada Dialogues

Agenda

Introduction of Sustainable Canada Dialogues and proposal for experimentation (Divya Sharma)
Presentation of Sustainable Neighbourhood Action Plan (SNAP) program (Sonya Meek)
Presentation of West Bolton SNAP (Jose Manuel Torcal)
Discussion of evaluation framework (all)

Sustainable Canada Dialogues Summary

In Fall 2016, and in the context of its Generation Energy initiative, Natural Resources Canada (NRCan) commissioned Sustainable Canada Dialogues (SCD) to examine how Canada could transition its energy systems to low-carbon configurations while remaining globally competitive. The scholars' resulting report, *Re-Energizing Canada: Pathways to a Low-Carbon Future*, offered specific recommendations on how to advance the low-carbon energy transition. One of the recommendations for immediate action was to support low-carbon energy transition experiments in innovative social practices and technologies covering the spectrum of diversity found in Canada. Transition experiments can be understood as deliberate interventions that explicitly embody a novel configuration of social and technical elements that could lead to substantial low-carbon change in energy systems. In May 2018, NRCan commissioned SCD to produce a 20-page report on a framework to evaluate low-carbon energy experiments, due in mid-September 2018. The focus on evaluation reflects the distinction between innovation and experimentation, where experiments are innovations that are explicitly evaluated to extract lessons learned to accelerate the low-carbon energy transition. On June 1, SCD thus invited a group of experts with relevant experience from the private, NGO and public sectors to Ottawa to a one-day working session to begin co-developing a holistic approach for evaluating socio-technical energy transition experiments and show how lessons learned can inform and accelerate Canada's low-carbon energy transition. To co-create the framework in a participatory process, the scholars will work with project proponents from four case studies, including the West Bolton Sustainable Neighbourhood Action Plan program.

Sustainable Neighbourhood Retrofit Action Plan Summary

The Sustainable Neighbourhood Retrofit Action Plan (SNAP) is a collaborative, neighbourhood-based solution for advancing urban renewal and climate action in older urban areas. SNAP helps municipalities improve efficiencies, draw strong local community support and build innovative partnerships for implementation of a broad range of initiatives in the public and private realms. By reframing environmental projects to incorporate greater social and economic outcomes, SNAPs help generate creative solutions that garner more support for implementation. Most importantly, SNAPs deliver

measurable results and scalable lessons toward the implementation of watershed plans and other municipal plans and strategies.

Individual SNAP Action Plans have four typical actions areas: (1) home retrofit programs that increase uptake in climate actions at home; (2) integrated infrastructure and public realm renewal projects; (3) integrated site projects and demonstrations on Multi-unit residential and ICI buildings; and (4) increasing community resilience and community connections.

SNAP is a multi-objective program, with broader sustainability objectives than just climate mitigation, including adaptation and resilience. The SNAP program evaluation follows four high-level methods for comparison: within neighbourhoods; between neighbourhoods; with mass marketing programs; and with control groups. The new West Bolton SNAP in particular will undergo an Interim Evaluation to assess change in behaviour, attitudes, and retrofit uptake as a result of the pilot Home Retrofit Program (early 2019), to evaluate the effectiveness of the Residential Home Retrofit Program, identify ways to improve the West Bolton SNAP Residential Home Retrofit Program, and inform similar programs in other SNAP neighbourhoods. Methods include intake and interim questionnaires, telephone interviews and control neighbourhood comparison—via a control residential survey and control interim survey. Additional evaluation activities are expected to occur after the Interim Evaluation by the SNAP Team to assess long-term outcomes of the project—for example, estimating GHG savings based on energy retrofit technologies or behaviours adopted.

Discussion Summary

The TRCA emphasised the utility of SCD, as a group of academics, in identifying best practices for evaluation and providing a robust framework for evaluation that the TRCA could use in its SNAP programs. The TRCA considers that third-party evaluation by academics can be more credible than self-reporting. While the TRCA's evaluation is designed to generate internal lessons and reporting to funders, a framework co-created with the academics could enable generating lessons to export the SNAP model elsewhere. Funding in general was mentioned as a key constraint on evaluation, in terms of being short-term and/or uncertain, and restricting evaluation results to being reported at the end of the fiscal year. Data accessibility was another constraint mentioned, as private utilities are unwilling or slow to provide data at the scale that is most useful to the TRCA. Resources in terms of personnel and capacity were also stated as limiting factors for a robust evaluation, as was the difficulty in attributing results directly to program interventions. The program's own adaptive nature means that any evaluation of its results must consider its continually shifting structure. It was emphasised that an evaluation would have to consider more than just low-carbon energy implications (e.g., resilience and climate adaptation) given the holistic sustainability-focused nature of the SNAP program. The TRCA felt that a mixed-methods (quantitative and qualitative) approach is most appropriate to capture the full spectrum of results. For example, the TRCA uses surveys, focus groups, follow-up phone calls, uptake rates and energy use maps to assess program impacts.

Discussion Minutes

SNAP Program Evaluation

The SNAP program has four high-level evaluation methods:

(1) Neighbourhood level. The TRCA tracks its participation and uptake and adoption of actions according to its own targets in each neighbourhood. Based on the funding they have, they commit to certain deliverables and track those, or they compare against long-term targets for that neighbourhood.

(2) Between neighbourhoods. The TRCA compares the typical rate of uptake and engagement between different neighbourhoods, despite having locally tailored approaches and program interventions.

(3) Program uptake vs. mass marketing. The TRCA compares its program uptake against mass-marketed programs for which it has access to data. Generally, they find their program uptake is up to orders-of-magnitude higher than the uptake of mass-marketed programs.

(4) Control neighbourhoods—Future evaluation. The TRCA would like to compare its neighbourhood uptake against control neighbourhoods that possess similar characteristics, in terms of demographics, building stock and sociocultural values. Taking like neighbourhoods and mindsets, how are uptake rates in their neighbourhoods, where they use a neighbourhood-based approach, compared to mass-marketed programs? Within its neighbourhoods, the TRCA would like to compare homeowner participants vs. non-participants.

Key Indicators

Number and type of intended retrofit actions

Number and type of retrofit actions taken by participants

Number of residents who have signed up to referred retrofit programs; change in awareness of these programs

Change in attitudes and perceptions towards retrofits

Participant satisfaction

Qualitative perspectives towards successes and needed improvements of the program

Number of residents who have signed up to referred retrofit programs compared to before program; change in awareness of these programs

They know specific actions implemented and have a GHG calculator, so they can estimate GHG savings

The value of SNAP is in trying to achieve other socioeconomic objectives beyond just sustainability and climate change. The other objectives, including climate adaptation and resilience, need to be included in any evaluation. By design, the program is multi-objective; the TRCA promotes what resonates first with homeowners and then gets them on a journey to adopt other actions. There are broader impacts beyond immediate home retrofits, on climate adaptation, resilience, promoting active transportation and electric vehicles, for example. Thus, the TRCA is interested in evaluating the holistic benefits that emerge from its program.

<u>Ideal scenario:</u>

Originally, the TRCA wanted to know the change in energy and water consumption, normalised for seasonality and weather conditions. However, these data are not accessible (see *Challenges* section).

Additional ideal indicator:

% of control neighbourhood residents that participated in utility retrofit incentive programs during the project period

Evaluation Process

What are the objectives? The reasons to evaluate/value of evaluating?

Externally: for reporting to funders on outcomes and to potential partners that the TRCA can attract and retain.

Internally: for continuous improvement of the program, asking: How can we be more cost-efficient? Do we need a new approach for reaching impacts? Are there elements that work that we should expand? How close are we to our targets? Where do we shift or adapt?

One objective is scalability: In this case, evaluation is to confirm whether the design for that neighbourhood typology is working and if the TRCA can scale the pilot to inform recommendations for scaling by themselves (to other SNAPs) or by others (e.g., other cities). The TRCA database lets them identify other neighbourhoods in which to promote their programs, by finding those with common socioeconomic values and overlaying these data on a map of building stock age.

For whom is the evaluation intended?

For funders—i.e. municipalities (who are core funders), provincial government (though not anymore) and federal government.

For program managers.

For the community—so the TRCA can show relevant learning or impacts and help them understand what the benefits are to them.

Who is carrying out the evaluation?

For evaluation to be seen as legitimate and valid by prospective funders or by other partners, it needs to be done by a third party, rather than as self-reporting. The TRCA longs for academic partners to lend more credibility; the benefit of peer review is to make sure the method is defensible and that the interpretations are validated by others.

The TRCA would want to explore the role of the project proponent in doing the evaluation vs. a third party. In West Bolton, they shifted to a more hands-on role in evaluation, whereas previously they were hoping to have it conducted mostly by a third party.

Who provides the funding for the evaluation?

Originally, the provincial Ministry of Environment and Climate Change (now the Ministry of the Environment, Conservation and Parks) was funding the evaluation portion of the program in West Bolton, but the funding has just been pulled following the change in provincial government. Apart from that, they use their own in-house time, primarily through municipal funding.

How are the results from the evaluation shared?

Via year-end progress reports to municipal partners, to support budget requests, and through final reports for grants.

Via conferences, webinars—local, provincial, national and international.

Via the SNAP program application to the C40 awards.

Verbally and informally with community members as a tactic to increase engagement. The program's 10-year anniversary will be an opportunity to report back on results with the community.

As part of the original funding, the idea was to create a knowledge-sharing network or "community of practice" to share lessons among behaviour-focused interventions that received money from the same grant. However, this will likely not continue, as that funding was pulled.

Ultimately, the TRCA's hope is to tell the story of SNAP and to encourage more people to do neighbourhood-based interventions and obtain more funding for them, and to inform other programs that they see could be made more effective.

How can the results from the evaluation be integrated in project implementation?

The TRCA will incorporate the results depending on what they are; if they are "good", the TRCA will continue doing what it is doing. If they show some opportunities for improvement, with targets left to achieve, then the TRCA will adapt the programs.

At the end of the year, when reporting back to partners, the TRCA takes stock and plans for the next year.

The West Bolton SNAP is the new SNAP the TRCA is designing now, thus they have the opportunity to update their research methodology and design and inform their baseline data collection to have a baseline against which to measure. This was not as consistent or well thought-out in the past. They can also track West Bolton over the next year.

What methods would/do you use? What data are available?

The TRCA will use mixed methods: The surveys will be hosted online, because they have an engagement platform for West Bolton. They will also conduct in-person, street surveys, and focus groups to get more in-depth understanding. In West Bolton, they will conduct baseline, ex-ante surveys for the first time. Lessons from West Bolton can be integrated into all the other neighbourhoods in the next generation of projects. The TRCA found a control neighbourhood within West Bolton that they felt comfortable with as a comparator.

The energy utilities or municipalities, in the case of municipal programs, have retrofit uptake and program participation data. They also have neighbourhood-level energy consumption data. However, it is hard to assess program success based on the latter, due to the presence of confounding variables like population growth, weather, etc.

Enbridge Gas and Toronto Hydro have offered to give the TRCA the Enbridge program uptake data for control neighbourhoods over the past six years. Thus, they are interested in comparing what is happening each year.

The TRCA is always adapting to see what works and doing more of that. Thus, a consideration in tracking and evaluating is that they might change the program each year. This means that evaluation can be quantitative up to a point, but the other side needs to be qualitative in taking rich insights from observations and making changes.

After the relationship with an academic soured (see section below), the TRCA is now working with a consultant to help them develop the framework and finalise it. Ultimately, they will deliver the survey, and the consultants will do the analysis, allowing them to maintain a third-party lens on it.

The TRCA was bound to the fiscal year-end reporting, but now that that funding is gone, they will push the follow-up survey to a later date when they will likely see more meaningful change. It usually takes 1–2 years from the time homeowners decide to implement an action to the time that it is actually implemented.

Ideal scenario:

The TRCA would have wanted true ex-ante baseline data gathering before entering the neighbourhood, since now they are dealing with potentially biased respondents. It is difficult to get an unbiased survey, because they first develop the plan in the neighbourhood itself. To do so, they would have to survey an area that does not have a SNAP at all. They would need resources (i.e., funding) to do that. Normally, their funding is focused on planning and not evaluation.

What have you learned as the program has evolved?

When the TRCA designed the Black Creek SNAP, the idea was to evaluate progress based just on neighbourhood energy use. They then decided to move to evaluate the uptake of actions, and not just calculate energy savings before and after the project.

The TRCA has experienced some issues with relationships souring with an academic who was to help them conduct the evaluation. A consultant has a contract that has legally binding deliverables, and so it is easier to maintain accountability than with academics. However, the value of an academic product could be a lot greater in terms of the ability to disseminate the results across the country. They have also had some great experiences with colleges on smaller research projects; they have found that the problem often comes down to scheduling constraints when term ends and students move on and the ethics review process.

Challenges to Evaluation

Data-sharing with utilities: Utilities take a long time to share data, citing privacy issues as the reason. Utilities are private companies—sharing data does not help their bottom line. When they do share data,

they will share by postal code, not by street. This means that it is so broad that it is hard to be able to relate it to what the TRCA has done in the neighbourhood. At one point, the TRCA had data sharing agreements and partnerships with the right people to support a smooth process of data sharing. But, historically, it has been hard to get data even from municipalities. There have been regulations in the past requiring them to share data. Since water utilities are public, they share everything—privacy does not seem to be a concern here.

Lack of sustained funding: The value of evaluation is not obvious in the absence of certainty around funding for the next year of the program. Even with sustained funding for the basic program, the TRCA receives year-to-year funding for specific areas or incentives of the program, and they do not know which incentives will be funded the following year. In an uncertain political climate, the availability of government funding programs also becomes uncertain.

Fiscal year reporting: The TRCA is constrained by reporting evaluation results by the end of the fiscal year, when that might not allow for enough time for meaningful results.

Timelines: Implementation of the project got ahead of the TRCA, and so they did not have time to design the evaluation beforehand as thoroughly as they would have liked.

Capacity limits: The TRCA is limited in resources and capacity. The more grants they get in a particular year, the more they can be out in the neighbourhoods and promote the program, and the more participation they are able to generate. There is more demand than they can fill; they could have higher numbers if they had more time to do consultations. Participation is directly linked to how much money they invest that year. Evaluation requires someone who is familiar with the area of practice, with behavioural approaches, community-based social market research and evaluative methodologies and who has experience in other jurisdictions with homeowner and home retrofit subjects—a specific skillset that is both necessary and rare.

Sample size: The sample size needed for high confidence results is not feasible, thus there will be no statistically significant results meeting a scientific standard. The level of effort to get 300 calls in a neighbourhood of 2000 homes is much too high. Statistically significant surveys, while not so relevant for the TRCA, lend credibility for peer review for any academic partner.

Adaptable program means hard to assess success: The SNAP program is successful, but it is hard to evaluate because it is adaptable—the basic structure stays the same, but every year there are slightly different incentives and marketing material that the TRCA adapts. Marketing for the initial, "low-hanging-fruit" residents might not work for the next cohort of residents. Thus, any evaluation needs to consider what marketing was used, which incentives were used and the level of staff they had (there is a direct relation between their effort and the outcomes).

Attribution of success: The TRCA will not know if there are other homeowners that would have been encouraged by them to retrofit but have not. This would have required household surveys to get that data rather than neighbourhood-wide energy use data. Some homeowners might have also retrofitted before their incentives were available. They could retroactively ask, but it is hard to attribute the action to their intervention.

Ideal Scenarios for Evaluation

The TRCA would like to be able to plan for a certain number of years in the future and have the freedom to have a multi-year program and multi-year evaluation strategy. This would allow them to follow the ups and downs, since it takes time to implement interventions. Normally, this is a challenge with the funding.

The TRCA could look at how they affect program uptake in the participating neighbourhood by surveying non-participants in the neighbourhood who did the retrofit themselves after seeing their flyers (i.e., there is more uptake than they can track) and then look for similar actions in non-participating neighbourhoods (i.e., Did they do it without the TRCA?). They could compare the uptake across neighbourhoods, and compare with experiences in other jurisdictions with programs of similar types. There should be different levels of evaluation within a neighbourhood, with those who have actively participated (the group for which they have good data because they call and follow up) and then the broader group (the whole neighbourhood, including those who were not approached). The TRCA would want to be able to look at this latter group separately too.

There is a theory that a program like SNAP has potentially higher levels of trust and openness coming from an NGO than as a similar program run by a municipal government. It would thus be interesting to evaluate the structural elements of the SNAP program in comparison with others, in terms of the delivery agent, to get causal explanations for success. As part of the neighbourhood design for each program, the TRCA researches who local residents trust and they have found that it is different in each neighbourhood: some trust the city, others trust the TRCA.

It would be interesting to ask: To what degree is the SNAP program making change within the municipal institution? How can the SNAP program help to spark change within the institution? What departments are connecting with each other that would not have otherwise? How are they improving alignment and coordination in a municipality?

The TRCA would like to have an evaluation that helps inform the particular messaging and framing of engagement materials. In the implementation process, the TRCA could test the rate of uptake across different forms of messaging (e.g., flyers vs. letters) in a single neighbourhood in the same year.

Role of Collaboration with Sustainable Canada Dialogues

The TRCA is interested in finding partners to help evaluate the SNAP model and the particular program interventions inherent within it. They cannot answer what they want out of an evaluation framework, what is implementable and what is ideal; they do not have a robust, ideal evaluation. The TRCA needs to know where to put the effort and what are the best practices. If SCD can help them come up with an evaluative approach and a framework, it would help them.

The TRCA is not an expert in how to evaluate, nor does it have the budget or access to what other jurisdictions are doing to which to compare its results. Instead, they are comparing the results against their own targets. The TRCA would like to know, for example, what is a reasonable pace of achievement for its targets, what is the experience elsewhere, and whether they are ahead of or behind the curve.

The TRCA has not put so much thought into how to export lessons externally—that is for groups like SCD to inform them how to do. For them, the interest is in continuous improvement of their own program. They do not necessarily have the capacity to engage nationally; while they did so with the

Federation of Canadian Municipalities, it has not been a main goal. Many programs report that their targets are fully met, but that is often because their targets are low, or they are only looking at a small proportion of the full market. Thus, there is a basis for the TRCA to develop toolkits for how they tailor their marketing and delivery to different neighbourhood typologies.

Appendix 4. NRCan Follow-Up Questions for Catherine Potvin and SCD to Guide Policy Discussions on the Proposed Energy Transition Experimentation Program

1. What is the evidence/rationale for developing an energy transition experimentation program?

The *low-carbon energy transition* entails not only the diffusion of new and existing technologies (e.g., the adoption of electric vehicles and new renewables) but also changes to the policy and regulatory frameworks (e.g., regulatory standards and carbon pricing regimes) to facilitate innovation and speed-up the adoption of novel social practices and norms (e.g., new perceptions of, and behaviors, around mobility and electricity use) (Geels and Schot 2007; Smith et al. 2005, Senbel et al. 2014). As clearly articulated by Wesley et al. (2011), resolving current environmental problems such as climate change will demand "harness[ing] human creativity and innovation potential to tip the interlinked social and ecological systems in the direction of greater resilience and sustainability."

In the context of necessary transformational changes, widespread and diverse experimentation is essential because we do not know exactly what the future will look like (Nilsson et al. 2011). There are many possible technologies (e.g., solar PV, solar thermal) and forms of social organization (e.g., cooperatives vs. private companies)—each with distinctive patterns of costs and benefits—that could move us toward a low-carbon future (Rosenbloom 2017). The characterization of the costs and benefits of different approaches, their distribution over time and space and their relative desirability and social acceptability cannot be fully known in advance. Denmark, for example, became the global wind energy leader following a "bricolage approach" in which owners-users shared knowledge and tested turbine designs (Etzion et al. 2017). Learning, therefore, serves as a core component of processes of system innovation, particularly the iterative relationship between 'doing' and 'learning' as transition experiments begin to carry forward transformative change on the ground (Loorbach and Rotmans 2010; Gross 2010).

2. What are the science/ knowledge/policy gaps that such a program would address?

Only by trying out different potential responses and novel energy modalities that help define low-carbon pathways can we begin to learn more about what these directions of change might entail. Yet, neither existing policy approaches (Sabel and Victor 2017) nor conventional management (Etzion et al. 2017) are well-equipped to confront the complex challenges described above. Governments thus need to move towards more adaptive governance (Potvin et al. 2017). However, while we know much about top-down policy, when it works and when it does not, there is a gap in implementing and rolling out adaptive governance. The OECD (2017) reports the need for governments to adopt orientations that enable appropriate risk-taking in policy and program design and delivery, and yet to remain tied to the principles of evidence-based policymaking. In that context, transition experiments have attracted considerable attention as a way to test alternative approaches systematically before locking in to new policies, and to take new and concrete steps to mitigate climate disruption while building knowledge, societal capacity to innovate and public engagement around the long-term low-carbon energy transition (Kivimaa et al. 2017, Hildén et al. 2017, Potvin et al. 2017).

3. What are the program's objectives and expected results?

The objectives of a transition program would be to accelerate and scale up the low-carbon energy transition across Canada by opening and energizing a space for experimentation, from which lessons learned can help promote deployment of appropriate and effective low-carbon alternatives, shift decisions made by consumers and develop evidence-based policy. In addition, it will be a showcase to the world of Canada's socio-technological approach to promoting a low-carbon energy transition.

Expected results from the transition program include:

(1) Collaboration with **hundreds of potentially transformative low-carbon energy transition experiments** defined as deliberate interventions led by innovators (e.g., provinces, Indigenous nations, on-the-ground businesses, municipalities, non-profit organizations) to instigate learning cycles from innovation design to adoption, implementation, evaluation, reflection, and to readjustment;

(2) The establishment and coordination of a **pan-Canadian transition learning network** facilitating an understanding of the multiple dimensions of scaling up. While simple scaling occurs where the intervention grows (but does not change) there are other scaling pathways that can be equally valuable—including modular (intervention copied and/or adapted elsewhere) and ecosystem (where an intervention creates new niches for other kinds of interventions) that could play a role in accelerating the low-carbon energy transition;

(3) The co-design of **policy on low-carbon energy system options**, with federal, Indigenous, provincial and municipal regulators, using evidence-based methods.

To meet these objectives, the transition program should:

- involve a comprehensive portfolio of existing and new transition experiments, covering different dimensions of the energy transition (e.g., housing, transport, energy efficiency and conservation, electricity production and consumption) in different provinces, territories, regions and cities, business sectors, etc., including experiments that address gaps in current programs.
- be structured not just for potential short-term gains (e.g., easily accessible emissions reductions or low costs) but for overcoming key bottlenecks, delivering co-benefits with other social and environmental priorities, and opening the door to further long-term change—while also reserving some share of the funding for proposals outside the strategic priorities, thus allowing space for genuine grassroots innovations. This could draw on a comprehensive literature review to identify effective experimental interventions internationally that have not yet been applied in Canada.[2] Overall, the transition program's goal is to unlock the transformative potential of innovation rather than optimize established trajectories.
- involve external advisory bodies to help design the transition program, establish a strategic framework and perhaps help adjudicate applications.

[2] Examples include the neighbourhood-scale solar experiment (Vauban) in Freiburg, Germany; the high-density district of Hammarby in Stockholm, Sweden, which incorporates circular metabolism and low-carbon development; and the BedZed sustainable suburban brownfield development in London, which is and managed by a partnership of public and private actors.

- mobilize researchers and technical experts to assist and evaluate the transition experiments and help systematize learnings and revise best practices.
- conduct systematic evaluation of transition experiments and lessons across experiments, using transparent metrics for comparative analysis and methods such as social science evaluation of experimentation outcomes.
- have a clear outreach dimension designed to engage and educate the public about the individual experiments, the transition program, and the low-carbon energy transition as a whole, and to communicate the lessons and experiences to wider audiences who can adopt, magnify, scale up and adjust relevant technologies, practices, business models and so on. This effort could build upon existing knowledge mobilization and outreach programs, under the umbrella of a broader 'social mobilization strategy' to guide and target engagement of Canadians on energy in all regions of the country, based on research evidence of "what works". Such a strategy should include youth education (curriculum and teaching tools) to inform and empower those who will be involved directly in the transition over the next 30 years.

4. What are potential qualitative/quantitative indicators to assess success?

Transition experiments can generate data, experiences, stories and examples to learn from, but they can also teach about processes—e.g., picking and doing the intervention, learning from it, etc. Transition experiments also have enormous educational potential as a citizen outreach and engagement tool while offering networking opportunities that can empower dynamic groups and individuals, allowing them to leverage their ideas, skills and energies to accelerate change. Thus, a range of indicators can be looked at. These will be highly dependent on the nature of the transition experiments undertaken and differ whether individual experiments or the transition program itself is evaluated.

We would favour a range of qualitative and quantitative measures appropriate to the complexity of the transition program and reflective of the wide range of trans-disciplinary expertise in our network. It will be important to go beyond simple measures, based on emissions or amounts of energy, to also look at scaling and entrenchment measures—indicators for how the experiments generate the conditions for scaling up (in multiple ways) and becoming sticky or hard to reverse (e.g., van der Ven et al. 2017). It is also important to consider synergies and trade-offs with other social and environmental priorities to offer a broader definition of success. A blend of surveys and/or interviews and focus groups might serve to gain in depth understanding of the experiments' legacy.

- The primary indicators of success at the transition program level will assess its contribution, within 2–5–10 years, to (1) help identify a pathway to the low-carbon energy transition, (2) develop the appropriate evidence-based policies to stimulate the energy transition based on lessons drawn from the successes and failures of each experiment, and (3) contribute to build up public support for the energy transition. We would ask: How many Canadians are moving towards low-carbon energy in regions where experiments are happening vs. in other regions? What are the perceptions of the people involved in these experiments? How many new innovations have been developed and integrated into experiments and are functional and adopted? Through these experiments, how much carbon emission is reduced from energy use in Canada? Economically, how many more companies and jobs are created and generate revenue for Canada?

- At the level of individual transition experiments, proponents will be asked to articulate a clear learning goal from their experience rather than simply designing and implementing a novel model. In support of this, there should be a plan to monitor and record project experiences using indicators that are both standardized or quantitative metrics such as GHG emission reductions, and also individual or more qualitative experiences about participant perceptions. Specific indicators can be developed drawing on existing frameworks of evaluation dimensions and indicators (e.g., the Pacific Institute for Climate Solutions evaluation report Table 3.3 (Sheppard et al. 2015)). Indicators should also seek to answer important questions such as: What are perceptions of the experiment within the host communities? How does being in a host community with an experiment shift community opinion or thinking regarding climate change and low-carbon transitions? Does hosting a transition experiment help provide a stronger sense of efficacy in shaping new social-ecological futures?

5. How would public confidence be obtained for this type of experimentation network?

In a program of transition experiments spread across the country that involve a diversity of stakeholders and project configurations, each experiment represents a point of contact, a glimpse of future possibilities and an educational opportunity for the communities, media, schools, local policymakers and various stakeholder groups. The experiments can be located in a particular place, supported by community values, and foster a local vision of the future that feeds into the broader low-carbon transition. As such, transition experiments will democratize innovation for sustainability and empower anyone who wants to participate.

To enhance public confidence, the transition program should be characterized by transparency, with results publicly available and accessible, as the basis for informing public discourse. Facilitating the aggregation of lessons could be supported through an interactive online portal where project proponents can access evaluations or team collaboration tool (e.g., https://slack.com) as well as record their findings. Data from the assessment and indicators previously described could be made widely accessible to encourage further knowledge creation through the broader involvement of adopters who can adjust and scale up relevant technologies, practices and business models. Such real-time knowledge-sharing can build pride, show change is happening on the ground and reinforce the notion that groups and individuals can make a difference.

In addition, the central involvement of researchers, using rigorous evaluation methods as a key element of the transition experiment network (see Questions 5 and below) would further build credibility, as polls show that academics are among the most trusted sectors of society (https://www.edelman.com/trust2017/trust-in-canada).

6. How would such a program fund differentiate from, or build-off of, existing energy transition experiments supported across Canada, including federal programs/initiatives?

Across Canada, a number of government and non-government initiatives are already underway that could serve as precedents or seed projects for transition experiments, including for instance: climate-change-related social mobilization projects such as the youth-driven "Do It in the Dark" challenge, which reduced energy consumption in student residences across university campuses (Senbel et al. 2014); novel developer-led building concepts such as the Active House, which simultaneously promotes energy efficiency and occupant well-being (Newman 2016); and the Centre for Interactive Research on Sustainability at the University of British Columbia, which was developed by researchers and the

university as a flagship for sustainable building practices and as a model of regenerative sustainability and living laboratory approaches (Fedoruk et al. 2015). At the provincial level, Québec's Fonds vert supports projects that, amongst others, tackle climate change, while Transition Énergétique Québec seeks to stimulate the low-carbon energy transition. At the federal level, NRCan's Low Carbon NRCan initiative aims at ensuring that the specific goals outlined in the Federal Sustainable Development Strategy are met. Both the Clean Technology Challenge and the Smart Cities Initiatives of the Impact Canada Fund are examples of existing programs on which the transition program can be built.

As proposed, a transition experimentation program would implement a systematic approach that aims to: support and give visibility to **alternative** energy transition initiatives; apply **comparative metrics** to evaluate those alternatives; and **iterate** on promising and partly successful initiatives and initiatives that are effective in other areas. Thus, the transition program could take advantage of existing programs financing transition initiatives by providing a way to evaluate and integrated them in the learning network. The transition program could thus could strengthen existing programs by:

- encouraging a more explicit understanding of such initiatives within the frame of low- carbon energy transition experiments;
- expanding and diversifying their scale and scope, e.g., by cost-effectively exploring additional alternatives to existing initiatives by municipal and regional governments;
- increasing communications and networking among different initiatives;
- focusing efforts and resources on key problems and bottlenecks that can open up wider movement;
- enhancing the mechanisms for systematic assessment, lesson drawing and sharing transition experiment experiences; and
- more systematic engagement with the public at all levels of government.

7. What weight would the program place on transition experiments that are social versus technological? Are there some early "winners" based on Canadian comparative advantages that should be pursued?

Instead of focussing on specific technological issues through prototype development, we envision a transition program of experimentation focused on the deployment of low-carbon energy transition at multiple scales, in the context of long-term sustainability and considering the specific social context of each location. Such transition experiments balance and integrate technological, business and social innovation and need to be evaluated on multiple dimensions. SCD's transdisciplinary expertise is a key advantage for engaging in such work.

While the relative emphasis of technological vs. social factors will vary from case to case, transition experiments that combine and integrate these dimensions will provide a more complete picture about evolving social innovation, practices and attitudes around new services and technologies. Experimentation should address ways to overcome key social barriers on energy advances and implement the 'behavioural wedge' as a key part of the strategy to meet carbon emission targets and build energy resilience. It should also serve as a way to nurture "innovation as a mean to transformation" (Westley et al. 2011).

Different provinces and regions of Canada might select certain areas for future energy development based on the specific know-how and energy resources available. In these conditions, it might be logical to support energy transition experiments that align with provincial/territorial development strategies. Municipal governments and utilities, especially those with ambitious energy or climate change targets, also often have ongoing programs that could efficiently be expanded and evaluated, e.g., the City of Vancouver's new Renewable City Action Plan. There are also examples of community-led and NGO-led efforts that are developing novel, locally appropriate and culturally-sensitive energy solutions (e.g., Prince Edward Island's sustainable community, Summerside; Drake Solar Landing community in Okotoks, Alberta; Toronto's Long Term Waste Management Strategy; T'Souke Nation solar project; and Tides Canada 'Project Neutral') that would benefit from evaluation and stronger profiling.

8. Are there obvious partnerships/co-funding relationships that should be, or need to be, part of a proposed program?

Energy transitions necessitate changes that will span societal actors, sectors and areas of expertise. A central element of transition experiments is that they forge partnerships among previously distinct groups in order to gain traction on climate change (Heiskanen et al. 2015). That is, "[e]xperiments typically bring together new networks of actors with knowledge, capabilities and resources, cooperating in a process of learning" (Berkhout et al. 2010, 262). In doing so, there is an opportunity to build lasting bridges among potential innovators (policymakers, change agents, entrepreneurs and citizens) as they collaborate in processes of co-creation and co-learning.

The transition program will thus engage with researchers from a range of disciplines such as social sciences, engineering, biology, economy, and more, but also with communities (including the support of local governments and citizens), private sector and local community organizations. Government partners will include federal and provincial ministries such as Environment and Climate Change as well as Industry and Infrastructure Canada, and more. Many possibilities for co-funding are available, such as the existing programs mentioned in Question 5. Of clear relevance to the transition program, the 2015 mandate letter of Canada's Treasury Board President stressed the importance of tackling "the real challenges we face as a country—from a struggling middle class to the threat of climate change", explaining that: "Responsible governments rely on sound evidence to make decisions to ensure we obtain good value for our money" and urging Minister Brison to "work with [his] colleagues to ensure that they are devoting a fixed percentage of program funds to experimenting with new approaches to existing problems and measuring the impact of their programs." In addition, as scholars, we are actively pursuing additional funding opportunities for the engagement of researchers with the federal program competition, *Network of Centres of Excellence*.

Regarding transition program management, we think that it would be simpler and more practical if the program were managed by NRCan. As mentioned above, many municipalities and regional governments are innovating on energy solutions in areas with strong energy/climate change mandates and are obvious partners.

9. What are the best practices in Canada, and internationally, to inform this proposal?

A number of lessons can be drawn from domestic and international engagement (Kemp et al. 2007) with transition experiments.

- First, transition experiments combine both top-down and bottom-up approaches to promote the low-carbon energy transition (Kivimaa et al. 2017; Sengers et al. 2016). That is, transition experiments link locally-enacted and societally-driven ideas with high-level priorities around climate change, visions for future low-carbon energy systems and appropriate resources—in the form of government funding and/or expertise, for instance. Third party actors, such as certain ENGOs, researchers and business associations, have also been instrumental in achieving energy solutions with communities (Sheppard et al. 2015).

- Second, there is an emphasis on different project scales and contexts, ranging from individual buildings or businesses to municipalities, regions and industries (van den Bosch and Rotmans 2008). This is consistent with the understanding that low-carbon energy transitions will need to be simultaneously enabled across differing scales and contexts (Brown and Vergragt 2008).

- Third, the most effective experiments link the attainment of climate- and energy-related objectives to the satisfaction of specific societal needs, including health and welfare benefits, improved service, functionality and convenience for consumers, comfort and cleanliness, enhanced community control or resource savings, sustainability and expanded commercial opportunities. The transition to a low-carbon society is about opening up new opportunities and meeting needs in innovative ways, not just about hitting mandated GHG abatement targets. Experiments also allow the exploration of different ways to improve communities, enhance welfare and build new business models while moving onto a low-carbon development trajectory.

- Fourth, it is important for transition experiments to avoid capture by incumbents who would either monopolize transition experiment funding or favour responses that do not necessarily align with more fundamental system transformation (Kemp and Loorbach 2005). Rather, it is the mobilization of ideas from societal innovators—in businesses, communities, public organization, or non-profits—that makes transition experiments a promising instrument (Loorbach 2007).

- Finally, international engagement around transition experiments suggests that there is as much to learn from failures as there is from successes (Romijn et al. 2010). Indeed, the notion of 'success' or 'failure' is not quite accurate nor as relevant for experiments, considering that each ultimately contributes to an open-ended and iterative process of co-learning and co-creation that promotes the scaling-up of new models to confront climate change (Bulkeley et al. 2015).

Considerations Towards a Program Proposal:

What is the role of the federal government versus SCD and/or other partners? What would be the role of Natural Resources Canada, Environment and Climate Change Canada, and other key departments in this program?

NRCan, on behalf of the federal government, would design the transition program and seek specific funding for it in Budget 2019. The funding would support (i) a portfolio of transition experiments across the country and in diverse sectors, (ii) the creation and operation of the learning network and, in Phase 3, (iii) the co-development of policy options.

It is possible to envision a transition program strongly directed from the centre with government officials and experts developing a list of experiments engaging with critical barriers to accelerated deployment of low-carbon technologies and practices. Alternatively, there could be a largely bottom-up process with

stakeholders driving the program by submitting transition experiment proposals for funding, an avenue favoured by SCD scholars.

Meanwhile, SCD would support the transition program objectives by (1) working with NRCan and with experiment proponents to develop a framework to evaluate the experiments and support the transition program, by deepening the understanding of each transition experiment, (2) identifying conditions that either facilitate or restrict implementation, allowing for efficient scaling-up, (3) promoting and taking an active role in the creation of a learning network among existing transition initiatives that will provide a context for the scholars to analyse results, synthesize experience, develop wider theoretical and practical insights and diffuse findings. NRCan scientists would be welcome to share these roles with SCD scholars.

Value of striking an arms-length advisory or governing body?

An arm's length advisory board could identify high-level priorities and innovation platforms in conjunction with locally- and societally-driven transition experiments. It could help foster public confidence and guarantee independence for the sake of objective evaluation of experiments, allowing for room to say what government cannot and offer "objective" critical assessment. It could, however, also result in lessening government interest in the transition program. Various possibilities exist; however, the final structure should be determined after careful discussions. Transparency and a structure allowing real-time engagement with stakeholders to stimulate cross-learning should be key elements dictating the governing structure. Regional advisory committees might also help adjudicate this transition program because of the diversity of regional energy systems, political economies and transition priorities, and the complexities of individual jurisdictions.

How should policy development address issues of:

- De-risking

As risks of failure increase with the degree of innovation, we stress the importance of policies designed to support testing, recognizing that some degree of failure is expected, and that knowledge gained from both successes and failures is put to good use. 'Surprises', a term proposed by Gross (2010) to replace failures, are valuable because they expose limits of our knowledge and areas of interest. If we plan experiments with the possibility of surprise in mind, we can learn from them and move forward, rather than be shut down by failure.

Transition experiments de-risk the innovation process in two primary ways. First, transition experiments may reduce risk for the private sector through partnerships and by pooling societal resources (e.g., leveraging public finances or lower interest rates available to public entities) and expertise (e.g., different skillsets and capacities). With this, they can incentivize riskier yet more transformative forms of entrepreneurial behavior among businesses (Burch et al. 2016). Second, transition experiments also enable appropriate risk-taking in the public service by reducing policy and political risk associated with public spending on innovation. They do so by diminishing the need to "know the answer in advance" or "pick winners" given that even transition experiments that do not deliver the expected outcomes can nonetheless deliver unexpected benefits, including considerable new knowledge that can contribute to an overarching learning trajectory and lead to long-term success (van den Heiligenberg et al. 2017).

- Partnerships

  Partnerships will be prevalent in the transition program at different scales. Partnerships will enable the implementation of the transition experiments *per se*; the different transition experiments will be *de facto* part of the program's learning network and, as such, new relationships between like-minded innovators and lessons will be shared. Finally, the transition program will seek partnerships with other networks to perform data meta-analysis and take advantage of ongoing innovations to extract lessons from early implementation. The transition program could formally connect smaller networks carrying out experiments within non-profit organizations (e.g., the Natural Step Canada's Energy Futures Lab), research institutes and universities (e.g., Sustainable Canada Dialogues), government (e.g., Energy Working Group, Natural Resources Canada), and industry associations (e.g., Combined Energy Options Ontario Working Group, Ontario Sustainable Energy Association).

- Learning

  The transition program will support a learning network that gathers together numerous initiatives that can then be analysed to draw lessons from commonalities and differences in pathways developed, external circumstances and outcomes. For example, policy engagement rests on a commitment to carefully assess results and reflect back on potential learnings. Ongoing and cumulative evaluation forms an essential part of transition experiments (Luederitz et al. 2016), generating "lessons that matter beyond the local context [and] aggregating lessons within and across sites" (Matschoss and Heiskanen 2017, 2). Assessment, in this context, relates not simply to direct transition experiment outcomes (e.g., the abatement of GHG emissions at particular costs) but rather involves the generation of new knowledge about how novel low-carbon models interact with existing systems (e.g., identifying and attenuating barriers) and could open up possibilities for the future (e.g., if the model were scaled up).

- Fostering of co-benefits

  Research on policy development will consider the fact that the low-carbon energy transition cannot serve solely to reinvigorate economic activity, modernize and exploit Canadian comparative advantages that matter in a carbon-constrained world. It should also improve citizens' overall quality of life and enhance justice and equity. Canada's transition to low-carbon energy systems will have to recognize historical, social and political injustices created by previous energy regimes, recognizing that low-carbon solutions are also new social practices that might conflict with traditional practices or cultural expectations and desires. Policy development must also consider unintended consequences of additional financial burdens on economically marginalized individuals.

Who would select the projects? What would the criteria be for such selection? Is there an early prioritization that would change over time vis-a-vis:

- large and small-scale opportunities?
- sector-specific vs cross-cutting?
- Cross regions; urban/rural?
- community-focussed and/or private sector driven?

These questions need to be elaborated within the next months. We consider these great questions and also underlined them in our Letter of Intent to the *Networks of Centres of Excellence* program. Clearly, the main sectors of GHG gas emissions in Canada should be tackled: transport, building and industrial processes. Collectively, scholars from SCD are interested in seeing actions at regional and local scales, in urban and rural settings, involving private sector and community groups while acknowledging the importance of working with northern and Indigenous communities. As indicated above, we therefore favour a transition program in which a large number of projects each receives modest support over large investments for a reduced number of initiatives.

Should an experimentation program have a phased approach (e.g., initial focus on specific sectors or social/technological areas; then build on early successes)?

In the sense that we recommend following best practices internationally and nationally, we favour a phased approach. Phase 1 (2019–2020) would draw early lessons from existing initiatives in Canada and abroad and establish the evaluation framework for transition experiments. Phase 2 (2020–2025) would focus on the implementation of new transition experiments largely developed from the bottom-up, while developing the learning network. Phase 3 (2025–2030) could possibly focus on scaling up the lessons learned in Phase 2. This phased approach would allow the transition program to operate with three "stock-taking" exercises five years apart, starting in 2020. These stock-taking exercises would align and feed directly into Canada's international knowledge-sharing at the Conference of the Parties of the United Nations Framework Convention on Climate Change.

Would increasing the number of supported projects (e.g., from 20-50) result in greater learning or impact? How would one balance program costing versus impact?

We envision a large number of transition experiments as an essential element to adequately drawing lessons. The Netherlands, for example, had hundreds of experiments (Kemp 2010)—we believe that is the scale for which we should be aiming. Experiments could also find more funding once the government injects money. One could imagine that we need transition experiments in different sectors (transport, industry, housing, urban planning, governance, business models). Given the diversity of assets and ways of life across Canada and the need to capture a variety of societal circumstances, having 5–10 transition experiments in each sector and in each province/territory would bring us to hundreds of transition experiments. Once the program is well-developed, the costing vs. impact could be evaluated after five years.

Are there any comparative experimentation programs (with similar objectives) or funds that exist globally?

Transition experiments have attracted considerable interest internationally (Hildén et al. 2017). Jurisdictions in Europe, in particular, have engaged in experimentation around low-carbon possibilities for several decades and there are hundreds, if not thousands, of specific transition experiments that have been carried out. Transition experiments have varied in terms of scale (regional vs. local), resources (funding), scope (buildings, electricity systems and transport), actor involvement (multi-sector partnerships vs. smaller groups) and context (from Europe to Asia).

One of the most prominent frameworks under which transition experiments were carried out was pioneered by the Netherlands (Kemp 2010). Between 2004 and 2007, the Netherlands expended 160.2 million euros to pursue roughly 150 experiments in transport, electricity and other areas under an energy transition agenda. Responding to climate change, energy security and resource depletion concerns, this jurisdiction deeply integrated transition thinking within policy and planning frameworks in order to drive low-carbon innovation across seven platforms: new gas, green resources, chain efficiency, sustainable electricity supply, sustainable mobility, built environment and energy-producing greenhouses. A directorate and taskforce were established to coordinate efforts, bridging the capacities and perspectives of different ministries and business sectors.

More recently, Finland has adopted a strategy to engender a culture of experimentation among government and societal actors (OECD 2017). This strategy aims to generate knowledge about the design of new legislative, institutional and operational models. As part of this, Finland has moved to create a platform of experimentation to address major societal challenges such as climate change and encourage innovation (Matschoss and Heiskanen 2017). A number of experimentation hubs are currently being developed and networked under the coordination of the Finnish Environment Institute, a government research institute created by the Ministry of the Environment. Under this framework, Helsinki (Finland's capital), for instance, is developing the smart district of Kalasatama. This district is envisioned as an experimental space and living laboratory through which to engage with major challenges and explore possible solutions. Experimentation is being supported at multiple scales, from mid to large scale projects around new forms of housing and smart grid applications to a portfolio of small-scale 'agile pilots' supporting citizen-driven innovation around environmental but also social and economic goals.

In addition to the government-led initiatives described above, it is important to realize that there are a large number of experiments out there, broadly construed. They tend to not think of themselves as experiments, however, and there are very few conscious attempts to experiment in the way that this implies—i.e., where a government undertakes the funding of explicit and open-ended experimental interventions in order to learn lessons, pick winners from among the experiments to scale up, etc. For example, in the United Kingdom the National Endowment for Science, Technology and Arts (NESTA) launched a contest with a prize of one million pounds, inviting communities to submit innovative plans for carbon reduction. This "Big Green Challenge" competition received over 100 projects for consideration (Westley et al. 2011). Several transition experiment projects and associated frameworks have also emerged in other contexts, including Asia (Bai et al. 2010; Berkhout et al. 2010; Seong et al. 2016). Taken together, a rich body of research has emerged around the deployment of transition

experiments across international jurisdictions, including documenting experiences and drawing lessons for future work (Hoffmann 2012; Kivimaa et al. 2017; Sengers et al. 2016).