Policy experimentation and policy learning in Canadian cultural policy

ABSTRACT
This article examines policy experimentation in the context of policy learning in Canadian cultural policy. Despite the attraction of experimentation to encourage learning and thus improved policy outcomes, much of the literature on experimentation does not give sufficient attention to how it is operationalized in practice. Drawing from a novel dataset based on interviews with key actors, this article focuses on how the governance of experimentation impacts learning resulting from experimentation. Findings ultimately demonstrate that while learning occurred, it was constrained overall by a hierarchical, top-down approach to experimentation. Lessons from this case study can therefore be useful for both policy scholars and public administrations embarking on experimentation or other types of public sector innovation in Canada and beyond.

1 INTRODUCTION

Shortly after Justin Trudeau was first elected Prime Minister in 2015, the Canadian government

unveiled a new governance approach focused on outcomes and public sector innovation. Based

on a justification that governments are facing ever-more complex challenges, while at the same

time "being asked to be more responsive to citizens, accomplish more with the same or fewer

resources, and be more accountable in administering public funds" (Privy Council Office 2019a),

part of this strategy mandated all federal departments to undertake experimentation. The goals of

experimentation were to *innovate* and *learn what works,* via testing hypotheses, in order to

improve policy and program outcomes (Privy Council Office 2016). This article investigates

experimentation in a department that embarked on a comprehensive experimentation program,

Patrimoine canadien/the Department of Canadian Heritage (PCH), responsible for federal

policies relating to the arts and culture, heritage, and sport.

There has been growing interest in policy experimentation by governments and public policy scholars, particularly in environmental and climate change policy, because of its potential to facilitate learning.[1] However, there has not been much attention paid to the *processes* of doing experimentation, which are contingent on governance arrangements and politics. Rocle and Salles (2018: 243) go so far to say that experimentation is a "form of politics." This concerns not only the questions that are asked in experiments (Brodkin and Kaufman 2000), but the way in which experimentation is operationalized. Learning is often a key goal of experimentation, but how does this actually work in practice? I argue that we need to understand the governance of experimentation to answer this question.

With this in mind, the article has two empirical objectives: firstly to assess the relationship between policy experimentation and policy learning, and secondly to analyze the governance of experimentation in this case study. This is accomplished via research methods – interviews with key policy actors – that offer insider views on how experimentation was conducted, alongside an theoretical framework that links experimentation and learning (Huitema et al. 2010). The case study of cultural policy is a 'least likely' example of experimentation in many ways, as experiments are not often associated with the arts and culture, and the outcomes of policies and programs in this field are difficult to measure (O'Brien 2010).[2] However, PCH's experience offers lessons in how to operationalize effective experimentation, since institutional rules can work to both constrain and enable learning (Heikkila and Gerlak 2013). It is particularly relevant

---

[1] Though see for example Brodkin and Kaufmann (2000) on experiments in social policy, Heilmann (2008) on economic growth and reform, and van der Heijden (2014) on the building sector.
[2] Cultural policies are generally agreed to be those associated with the arts (museums, visual and performing arts, heritage, and literature and poetry), and may also extend to other areas such as sport, languages, libraries, zoos, botanical gardens, fairs and festivals, folklore, and crafts (Mulcahy 2006).

for Westminster systems where there is often a low tolerance to risk-taking (Aucoin, Smith, and Dinsdale 2004). Lessons from this case study can therefore be useful for other jurisdictions embarking on experimentation or other types of public sector innovation. The article demonstrates that the potential of experimentation will be under-realized unless it is properly resourced, and the tension between the reflexivity required in learning and the hierarchical approach of monitoring is resolved.

I begin in Section 2 by locating this research in relation to existing literature on experimentation and policy learning. Section 3 contextualizes experimentation in the Government of Canada and more specifically in PCH. Section 4 then discusses the research methods of document analysis and interviews. I then outline the findings in two sections, showcasing the key factors important in governing experimentation in PCH as well as what kinds of learning occurred and what lessons can be drawn from this case study.

2 POLICY EXPERIMENTATION AND POLICY LEARNING

The first mentions of experimentation in policy can be traced back to seminal works by scholars such as Donald Campbell (1969) and John Dewey (1999[1927]), who argued that all policy-making is experimental in some sense, as no one can completely predict outcomes. However, the literature on policy experimentation has developed in a piecemeal fashion.[3] The concepts of *experiment* and *experimentation* are "...often used casually and with great rhetorical effect" (Ansell and Bartenberger 2017: 36), and no single definition of experiment exists (Hildén et al.

---

[3] A keyword search of 'experimentation' and 'public policy' into Web of Science shows experiment used in a variety of ways, including an experimental approach to governance (Sabel and Zeitlin 2010); to describe scientific work undertaken exogenously to government, the results of which are then used to inform policy-making (Stoker 2010); and as a tool for learning to inform policy-making (McFadgen and Huitema 2017), to use just three examples.

2017). In general, there are two main understandings (Ansell and Bartenberger 2016). The first is experimentation as randomized control trials, where interventions are tested in a scientific, rigourous manner. The second understanding more broadly conceptualizes experimentation as an innovative, new way to solve a problem. Ultimately, experimentation is a "purposeful and coordinated activity geared to producing novel policy options" (Heilmann 2008: 3) in which taking small, calculated risks can inform future decision-making and improve outcomes.

Learning is a key element of experimentation, as "the intention to learn is at the core of what it means to experiment" (Ansell and Bartenberger 2017: 38). Learning is "the process of updating beliefs about public policy" (Dunlop and Radaelli 2018: 256). It "arises from dialogue about alternatives and reflection about alternative interpretations of evidence" (Mahler 1997: 523). It is also "eminently (albeit not exclusively) about power" (Radaelli 2003: 13), in that processes of deliberation operate according to particular logics and values, structured by a system of rules (Dunlop and Radaelli 2018). Learning is analyzed herein as an emergent process, produced via socialization. It is therefore dependent on relationships, conditions, and structural factors (Heikkila and Gerlak 2013; Sanders et al. 2020).

The literature on learning in collective settings points to several contextual factors that enable learning. Heikkila and Gerlak (2013) discuss three main categories: institutional structures, social dynamics, and technological and functional domains. They show how factors such as the organization and coordination of tasks and responsibilities, communication patterns between actors, and the resources and technical systems available to actors, can both enable or constrain learning. Dunlop and Radaelli (2018), meanwhile, discuss how different decision-making

contexts influence what kind of learning takes place, and how. For example, they show how learning in hierarchy is difficult due to the emphasis on rule-following and compliance, which discourages reflexivity and innovation.

Similarly, the extant literature on experimentation points toward four key factors important for enabling experimentation.[4] First of all, multiple studies point to the importance of *leadership* that is legitimate, supportive, innovative, and committed (Farrelly and Brown 2011; Bos et al. 2013; McFadgen and Huitema 2017; Rocle and Salles 2018; Brown and Cohen 2019; McFadgen 2019; Suškevičs et al. 2019; van Doren et al. 2020). Secondly, clearly-defined and communicated *resources* (Bos et al. 2013; Newig et al. 2016; Witting 2017), transparent rules (Witting 2017), and flexible administrative practices (Bos et al. 2013; Heiskanen et al. 2017) are more likely to lead to effective experimentation. A third category is broader *cultural factors*. These included the importance of a sense of ownership, persistence, and dedication by those carrying out the experiments (Farrelly and Brown 2011; McFadgen 2019; van Doren et al. 2020), a curious and adaptable attitude (Farrelly and Brown 2011; Brown and Cohen 2019), internal motivation for learning (Suškevičs et al. 2019), trust, and clear, participatory communication (Farrelly and Brown 2011; Witting 2017; Suškevičs et al. 2019; Sanders et al. 2020). Where stakeholders are involved, buy-in is also imperative (Farrelly and Brown 2011). Finally, on *evaluation*, the literature suggests that appropriate resources for evaluation are crucial (van Doren et al. 2020) and that a light-touch approach to evaluation can encourage reflexivity (Brown and Cohen 2019).

---

[4] For this literature review, I first searched "policy experiments" on Web of Science. I discounted sources that did not address the governance of experiments. I then made a list of factors that were important to the realization of experiments, and loosely grouped them according to the four themes.

Based on this review, it is reasonable to expect that 'learning what works' is more likely when these factors are in place, enabling a setting that encourages reflexivity and flexibility.

There is no agreed-upon definition of a successful experiment in the literature, nor in the policy documents produced by the federal government or PCH. Success could mean simply performing an experiment, or conducting a good-quality experiment, or that an experiment's hypothesis proves true. It could also mean, more broadly, the implementation of an experimentalist approach to policy and governance. However, because learning is at the core of experimentation both theoretically (Ansell and Bartenberger 2016, McFadgen and Huitema 2017) and in the Government of Canada's stated objectives of the approach (Privy Council Office 2016), the analytical focus of this article is on learning: did learning take place, and if so, what was the nature of that learning?[5]

The relationship between learning, experimentation, and the governance of experimentation is not empirically well-established. However, a central contestation of this article is that *effective experimentation involves learning*: experimentation involves identifying a public policy problem, positing a change that might lead to improvement, and carrying out the experiment to learn whether that has worked. Experiments produce knowledge which can then be used for political decision-making (McFadgen and Huitema 2017). Those who experiment are setting out, at the most basic level, to learn something. Learning occurs both when a correct hypothesis has been reached (positive result) and when it has not been reached (negative result).

---

[5] It should be noted that most public servants did not conceptualize success this way; for them, it was mostly about completing a task. Thanks to an anonymous reviewer for drawing my attention to this.

To conceptualize and analyze learning, I draw on a typology developed by Huitema et al. (2010) and others (see for example Baird et al. 2014 and McFadgen and Huitema 2017) that identifies three types of learning: cognitive, normative, and relational. Cognitive learning refers to the acquisition of knowledge, in order to improve existing knowledge. This could be, for example, the discovery of previously unknown effects via experimentation (McFadgen and Huitema 2017). The second and third types represent changes as a result of the acquisition of new knowledge. Normative learning is a change in perspectives, norms, values, or "convergence of group opinion" (Baird et al. 2014: 53), which may then lead to a change in policy or program. Relational learning leads to an augmented understanding of others' mind-sets, increased trust and cooperation, and the building of relationships. Normative and relational learning are therefore deeper types of learning that involves changes in views, opinions, and understanding.

< Table 1 >

While this typology has been primarily used to assess learning from experiments in environment and climate change policy, it can also be applied to those in cultural policy. Indeed, the typology's creators encourage its application in other policy fields (McFadgen and Huitema 2017). It specifically allows for the identification of learning that occurs as a result of the direct outcomes of doing experiments (rather than other types of knowledge acquisition such as individuals seeking information from other jurisdictions). It also allows for the identification of outcomes of learning that occur when working with external stakeholders, a feature of many of PCH's experiments. Thirdly, this typology accounts for collective learning. Understanding the link between individual and collective learning is a theoretical challenge in the policy learning

literature (Heikkila and Gerlak 2013). In this case, conducting experiments was a team activity, involving acquiring, assessing, and disseminating information, via discussion and deliberation among team members (Heikkila and Gerlack 2013).

3 CASE STUDY: CULTURAL POLICY AND THE DEPARTMENT OF CANADIAN HERITAGE

In his October 2015 mandate letter to then Treasury Board President Scott Brison, Prime Minister Trudeau explained that experimentation would involve trying new things, and measuring the impacts of new approaches (Trudeau 2015). This would mean "a strengthened culture of measurement, evaluation, and innovation in program and policy design and delivery" (Trudeau 2015), reflecting an increased emphasis on evidence-driven, results-based outcomes.[6] This approach was influenced by Michael Barber's idea of *deliverology*, which focuses on outcomes and results (Barber 2015; see also Birch and Jacob 2019).

Over the next year experimentation was clarified and more detail was provided by the government's Impact and Innovation Unit (part of the Privy Council Office), as well as the Treasury Board Secretariat. Experimentation was defined as "testing new approaches to learn what works and what does not work using a rigorous method"; the aim was to use innovation to "find new ways to address persistent problems that traditional approaches have failed to solve" (Privy Council Office 2016). Examples of methods to do so included "deliberate, thoughtful, and

---

[6] The emphasis on evidence-based results and outcomes, as well as increased monitoring and reporting, is indicative of neoliberal approach to public sector innovation more broadly. Space constrains this discussion in this article, but future work can and should expand on this (see also Birch and Jacob 2019).

ethical experimental design," comparisons such as randomized controlled trials, A/B testing, and rigorous impact measurement (Privy Council Office 2016).

Patrimoine canadien/the Department of Canadian Heritage (PCH) is the federal department responsible for creativity, arts, and culture; heritage and celebration; sport; diversity and inclusion; and official languages (PCH 2020). PCH is a relatively small department with just over 1,800 employees (Treasury Board of Canada Secretariat 2019). It is a grants and contributions department which means that the vast majority of its budget is devoted to funding programs (PCH 2018a), many of which are delivered through the department's five regional offices across Canada.

PCH itself recognizes that "[a]t first glance, the Department of Canadian Heritage might not seem like the type of organization that would be highly conducive to experimentation" (PCH 2018b: 9). It is not a science-based department such as Health or Environment and Climate Change. Moreover, PCH has "many long-standing programs and long-term funding recipients," leading to "an environment where some stakeholders prefer the status quo" (PCH 2018b: 9). In addition, given the challenges of determining the impacts of arts and cultural participation (O'Brien 2010), the outcomes of its activities are not straightforward to 'measure.'

Despite these challenges, in 2016-2017, the department embarked on an ambitious experimentation plan in which every program created an experiment. There were 45 experiments (PCH 2018b).[7] 42 of these related to funding programs, and three were associated with corporate

---

[7] In the end, PCH had 51 experiments, but six were created after the publication of the plan.

services. The former were created and run by program teams. Generally, the aims of experiments fell into three categories: administrative efficiencies, improving access to PCH's programs, and capacity-building. Experiments with goals related to administrative efficiencies aimed to administer the program in a way that was more effective for the sector/applicants and/or the program team, such as testing new survey designs to get feedback on programs. Secondly, a number of experiments focused on expanding eligibility and reaching new recipients, particularly with Canada's Indigenous and ethnocultural communities. The third impetus for experiments correlated directly with improving capacity in the sectors that PCH serves, such as supporting new types of collaboration, trying to learn something about how the sector worked, or improving the sector's access to information.[8]

In experimentation, risks were allowed in a measured, calculated way, a seeming departure from a general setting usually characterized as risk averse (Aucoin, Smith, and Dinsdale 2004). Some failure, a concept not often associated with a government's goals (and particularly the goal of improving outcomes), was permitted: PCH emphasized that they expected 10-30% of experiments to fail, either by achieving a negative result hypothesis or by not following through on an experiment (PCH 2018: 22). Achieving correct hypotheses across the board, they argued, meant that the "level of ambition…is too low" (PCH 2018: 22).

4 RESEARCH METHODS

In order to determine how experimentation was operationalized and whether it led to learning, insider information from key actors was necessary. I carried out a two-step process. The first step

---

[8] A full outline of the department's experiments can be found in the Report on Experimentation (PCH 2018b).

was a detailed review of key policy documents, all of which were publicly-available.[9] In the

second stage, I carried out semi-structured interviews with civil servants based in the Department

of Canadian Heritage (PCH), Privy Council Office (PCO), Treasury Board Secretariat (TBS),

and Innovation, Science, and Economic Development Canada (ISED). Interviews with

individuals in PCH allowed me to understand experimentation as it applied to my case study,

whereas engagement with PCO, TBS, and ISED gave me a broader understanding of the general

context of experimentation across the Government of Canada. All interviews took place between

June-August 2019.

In total I spoke to 30 individuals across 18 interviews, representing 16/29 (55%) program and/or

policy groups and 23/45 (51.1%) experiments listed in the report.[10] Interviewees represented all

four sectors that had experiments in the department. They included members of program teams

that had both hypotheses that proved true (positive hypotheses), those that did not (negative), and

those who had created an experiment but did not go through with it. They also included

individuals who worked in corporate services and had knowledge of experiments and

responsibilities for overseeing them, but did not run any themselves. Most interviewees were

manager level, though I spoke to a range of individuals at varying levels of seniority. Fifteen out

of 18 interviews were in person and three were by phone. All interviews were anonymous.

Interviewees are identified herein by a letter or a letter/number combination.

---

[9] These included the 2015 mandate letter from Trudeau to Brison; the 2016 Privy Council directive on experimentation (Privy Council 2016); the Impact and Innovation unit's annual reports; and the Experimentation Works (2019) blog. In PCH, key documents included the 2018 Report on Experimentation (PCH 2018b); the 2018-2019 department report (PCH 2018a); and the department's evaluation plan (PCH 2018c).
[10] Using the report on experimentation (PCH 2018b), I emailed an interview invite to all programs that had an identifiable team member listed in the government's Electronic Services Directory of Public Servants. Of the programs I did not reach, some did not respond to requests for an interview, and for others I could not find contact information.

To analyze the material, I undertook a three-step process. I first read through all documents and interview transcripts and notes, in order to form a general picture. In step two, I analyzed the data more closely, interview-by-interview, picking out broad themes relating to the governance of experimentation and processes of learning. I also began to identify types of learning. Evidence of learning came from both the experimentation plan, which detailed the results of the few experiments that had been completed at the time, as well as my conversations with civil servants. For each experiment, I analyzed the problem identified, the interventions, the desired results and hypothesis, and, if the experiment was complete, the results. From this data, I determined whether new knowledge had been acquired (cognitive learning), and then whether or not that had led to changes in values or opinions within the program team (normative learning) or whether, for example, relationships or cooperation had been deepened, usually in the context of stakeholders, as in relational learning. An important aspect of this was identifying the conditions in which different types of learning emerge. In a final review of the data, I refined the four themes below more comprehensively and chose illustrative examples of learning. The article therefore draws together both how experimentation was governed, how learning occurred, and how the two are related.

5 FINDINGS: GOVERNING EXPERIMENTATION

In this section I focus on understanding how experimentation was operationalized, broken down into four subcategories: leadership; resources and support; reporting and evaluation; and cultural factors, encompassing stakeholders and attitudes to risk.

Leadership and the top-down nature of experimentation

The broader context of innovation and improved policy and program outcomes was a political priority of the new Liberal government in 2015. However, departments engaged with experimentation in a very mixed fashion. In PCH it was spearheaded by the former Deputy Minister (DM), a member of a government-wide task force on experimentation, who was largely responsible for driving experimentation in the department.

One key aspect of the way that experimentation was operationalized in PCH is that it was mandatory: programs were *directed* to experiment. This created some uncertainty and confusion. Some civil servants felt that it was first presented to them in a way that was not immediately accessible:

> "... we all felt 'this is pretty cool,' but they were giving us examples of experiments at Health Canada, for example, in a laboratory, with some pills you know, and we're like looking at each other like 'how does this apply to us?'" (Interview C).

Moreover, many felt that it was a methodology-first scenario, where teams started out with a directive to do something, rather than a problem to solve. This forced nature was a barrier to overcome for many programs:

> "People don't want to make up problems just to say that they have an experiment" (Interview R)
> "You can't force someone to be creative - that's not how it works" (Interview T)

Some civil servants felt that the framework for experimentation was "very rigid, […] not as freeform as it should be" (Interview Y); that the definition of an experiment was too narrow and prescribed. At least one manager said that their team had many ideas, but because they did not

13

have a hypothesis and target group, they could not be counted as experiments and did not go ahead.

Experimentation, though Government-wide, was perceived in PCH to be tightly associated with one individual, the then Deputy Minister. Because of this, some felt that it was at risk of being a 'flavour of the month' and then disappearing, raising concerns around continuity and institutionalization - "making it part of the culture" (Interview P). At the time of fieldwork in summer 2019, some interviewees felt that it was declining in priority, as the former DM had left and moved to another department. There was thus a reluctance among some individuals and teams to fully engage in experimentation because of this feeling that it would not last. A number of interviewees mentioned this particularly when it came to the sustainability of budgets and the ability to keep adjusting their programs: if the funds for experiments came from a specific pot of money, what happens when that is gone?

Resources and support

Because experimentation was a new initiative, there was both government-wide and department-specific support put in place (see Brock 2021, Experimentation Works 2019). Alongside centralized oversight in the PCO's Impact and Innovation Unit, a new team at PCH was created in part to support experimentation. The Treasury Board also implemented a new procedure for programs that wished to make changes to their terms and conditions (rules about how public monies can be spent). This process involved six financial instruments that were meant to expedite what would otherwise be a complex and long process (see PCH 2018b: 13). Despite this, there were varying opinions on whether support for experimentation was sufficient. Some

felt that change was easier because fewer allies and senior level approval were required. Many managers also felt well-supported by their director. However, a number of interviewees also reported that running the experiment felt like an 'add on' to their normal job, that it was never their number one priority: "...you're trying to keep the lights on. That doesn't give a lot of time to do this additional thinking that experimentation or innovation requires" (Interview R).

Despite acknowledging that the "department is doing its best to organize its internal supports [...] [to] help to nurture experimentation" (PCH 2018b: 18), there were challenges relating to support and resources. For example, the Innovation and Experimentation team started *after* experimentation was already underway in PCH, making it more difficult for them to support experiments that had already been approved. This team's mandate was to do light reporting and bring other innovative methods such as design thinking to PCH, but it was also under-resourced, consisting of only two people for the first six months. Other constraints included corporate services bottlenecks in areas such as human resources, procurement, and risk management. One manager explained that it had taken them a full year to modify the terms and conditions of their program, which delayed their experiment. Others mentioned *wanting* to use the terms and conditions add-on tools from the Treasury Board, but being met with "huge resistance" from the Treasury Board itself: "even though there were those flexibilities that were put forward, and we tried to use those tools, it wasn't as flexible as we assumed it would be" (Interview X). Overall there was a feeling among many that the drive to 'do different' had not trickled to corporate services; instead, corporate services were faced with an increased volume of work without proper support to, in turn, support experimentation.

In terms of professional development, teams appreciated time and space to think about things differently. Several people felt that experimentation was gratifying and that they grew professionally from having undertaken it. Some managers felt they had gained new skills and ways of thinking, indicating that if they were to move to another program or department they would bring these skills with them. Particularly interesting were those who reported a greater sense of agency – the idea that anyone could put an idea forward as a basis for an experiment. In more than one case managers admitted to being pushed to act against their own biases: "it helped me as a manager to undertake something that my own bias might have delayed me doing," which "developed [them] as an employee" (Interview H).

Reporting and evaluation

In general, there were two main concerns raised with regards to reporting and evaluation. The first was that the reporting relating to experimentation was time-consuming. Many managers mentioned that they still had to report in a 'normal' way: "bureaucracy took over and killed it" (Interview T). Evidence from interviews suggested that people wanted a less rigorous approach to reporting on experiments.[11]

The second challenge was more general. There was a feeling that outcomes that went beyond quantitative data were not easily captured or taken as meaningful by management:

> "it was very easy to generate good-looking dashboards of progress and charts and graphs, but they don't capture all the meaningful sorts of outcomes" (Interview Z).

---

[11] "Reporting" for most interviewees consisted of all administrative work outside of the experiment itself, such as documents for TBS, surveys from the Innovation and Experimentation Team (twice per year), as well as Management Accountability Framework reporting, an annual government-wide assessment of management practices and performance (Treasury Board Secretariat 2016).

One interviewee described how they wanted to do a particular experiment but because it would not result in quantifiable results, it did not go ahead, as "for our program you always have measurable results at the end of a project" (Interview C). Another described it as, "just because something can't be proved [in an experiment], it doesn't mean that the approach doesn't have value, that we shouldn't try to improve things" (Interview A2). There were also difficulties for experiments that were trying to engender long-term change. One manager felt that in their program, for example, change would only be noticeable in 10-20 years because of the nature of the experiment and sector.

Overall, many expressed disappointment that the resources did not exist to take a different approach to reporting and evaluating experiments. While experimentation was meant to encourage imaginative, decentralized thinking among program teams, many felt that this was stymied by the obligation to measure and track in a top-down, centralized fashion.

Cultural factors: stakeholders and attitudes to risk

In interviewing PCH employees, I found a group of people who were highly committed to their work and worked diligently to support heritage and culture across the country. Some civil servants were artists themselves, and passionate about the arts and culture. Job satisfaction is high in PCH in relation to the rest of the public service: in 2019, 86% of PCH employees responded that they strongly agree or somewhat agree that they are satisfied with their department, compared to 71% across the public service (Canada 2020).

There was also a sense of a culture of risk-adversity and preference for the status quo, something the department acknowledges in their report on experimentation (PCH 2018b). One of the reasons for this is that the department has many long-term funding recipients. In experimentation, communicating and negotiating experimentation with stakeholders had mixed results. Many program managers explained that they had close relationships with the sectors that their program supported. I found that in some cases experimentation allowed for greater collaboration with stakeholders. In at least one case the experiment had gone so well that one manager reported that the recipients of additional funding created by their experiment said "you have no idea what this means to us" (Interview C). Another manager said that their experiment had been a "resounding success" (Interview P) for the stakeholders.

However, for some, changes to programs in the form of experiments meant tricky conversations, with some equating it to 'damage control.' Changing resource allocation meant that some recipients were going to gain and some lose, reflecting the contentious and political nature of experimentation (Brodkin and Kaufman 2000). The question became how to communicate and frame the changes. Some managers had to tread very carefully in these conversations: "We're very oversubscribed, so to suddenly talk about carving out some of those oversubscribed dollars, for a harder-to-explain, speculative purpose, is very tricky" (Interview Z). They commented that because there was no overall communication from PCH to program recipients, stakeholders, or the broader sectors, this made their jobs more challenging. An individual outside of PCH commented that specific to Heritage was the "reluctance to mess around with stakeholders" (Interview E2), and linked it to little incentive to change and damage often well-established relationships.

A second important cultural factor was attitudes to risk, a common theme in conversations with interviewees. The perception of latitude, and of changing attitudes to risk, differed among interviewees. Many said that they appreciated that experimentation encouraged them to be bolder and to take risks, where before it only occurred "on an ad hoc basis" (Interview D). There was space to do more without an end goal, particularly when compared to program development. Some thought there had been a noticeable change in attitudes:

> "...it does force us to think outside the box and to look at things differently, not just manage your programs in the same fashion year after year" (Interview X)

> "We have more discussion about bold ideas. That's a big difference" (Interview B)

Other interviewees felt that experimentation had encouraged more strategic thinking across the department by the mere fact of formalizing it. One manager described it as a further push to ask "'are we even asking the right questions?'" (Interview E).

However, many commented that even though experimentation promoted an atmosphere of dynamism, there were also constraints related to a low tolerance to risk, entrenched in the political culture in the form of ministerial responsibility and parliamentary oversight (Aucoin, Smith, and Dinsdale 2004). Many interviewees described the ease of keeping the status quo, even when it is mediocre, and the many challenges of achieving change in government: "everything" is aligned against the idea of experimentation in the public sector (Interview E2), according to one interviewee. Reflecting on a long career in government, as another interviewee put it, "we had more freedom 20 years ago. It's getting unfortunately more difficult to actually try something. [...] we still had to colour within the lines" (Interview Y). However, some also commented on the risk of *not* taking risks, demonstrating that this thinking is not ubiquitous:

"...there may be opportunity costs to doing the same thing until the end of time! If you never change your approach, you don't know what risks you're taking - what opportunities you're failing to make the most of" (Interview Z).

6 FINDINGS: IDENTIFYING AND EXPLAINING LEARNING WITHIN EXPERIMENTATION

This section concentrates on what learning occurred in PCH's experience with experimentation, what factors led to learning, and what lessons can be drawn from this case study. This discussion should be seen as an early analysis of the impacts of experimentation in PCH, since at the time of fieldwork, many experiments were still ongoing. However, this allowed for a focus on the processes of doing experimentation at various stages. I show that while learning did occur, it was underdeveloped due to the top-down, hierarchical nature of experimentation and a lack of resources.

6.1 What learning occurred?

I found evidence of all three types of learning as identified in the typology. Of the 23 experiments, eight were still in progress. Though I discussed with those teams what they anticipated learning from the experiments, the type of learning is not identified in the table below, due to the possibility of unexpected results. However, I do share examples of anticipated learning in the discussion below. In one case (Experiment 5), I did not gather enough evidence to determine the outcome of learning.

Cognitive learning, an increase in knowledge, was the most common type, occurring in nine experiments. For example, program teams learned about features of the sectors they funded, such as what was important for spatial planning for the future of the sector. In another example, the program team experimented with new ways of collecting feedback from program participants (Experiment 11). They tried a number of approaches and found the one that resulted in the highest response level. Another team experimented with different ways of marketing a particular cultural product (Experiment 19) and learned that "we're still not doing enough to communicate what we have to offer" (Interview M).

< Table 2 here >

Normative and relational learning were much less common in the evidence to date. The case of normative learning (Experiment 2) resulted from initial cognitive learning, whereby the testing of one hypothesis led to a change in perspectives. The program team hypothesized that if they changed a specific program requirement, they would see more uptake from Indigenous applicants. They did, and the outcome was as they expected. This change has been permanently adopted in the program. There were also anticipated cases of normative learning in experiments that were still ongoing. This was most evident in programs that tried to increase engagement with Canada's Indigenous and ethnocultural populations, who have been underserved in an arts and cultural landscape that privileges Eurocentric forms of culture (Paquette, Beauregard, and Gunter 2017). Normative, or anticipated normative, learning was more common among teams that had discussions about the overall aims and goals of their programs and how Canadians might be better served.

In the case of relational learning (Experiment 9), the program team developed new relationships and understandings of labour conditions in the sector, via direct engagement with creators as a part of their experiment. In this case, the program learned how to better communicate with creators as well as a deeper understanding of employment in the sector. This team also regularly worked collaboratively with another federal department, which they said helped them learn as they were used to openly communicating with actors beyond PCH. In another example that was still in progress and so is an *anticipated* case of relational learning (Experiment 21), a program manager described that their experiment had led to meaningful, ongoing dialogue with stakeholders. This particular experiment was co-created with stakeholders, meaning that the program team gained increased understanding of their views. This evidence suggests that open communication and stakeholder buy-in increased the likelihood of relational learning (McFadgen and Huitema 2017).

Overall, then, experimentation did lead to learning. Learning occurred when hypotheses were proven true and when they were false. Programs with experiments that did not go ahead – a "less helpful" type of failure, according to the department (PCH 2018b: 22) – learned little about their experiment but may have gained skills in experimental methodologies. However, the depth of learning was in most cases quite shallow. This partly had to do with the scope of experiments; an experiment that was about administrative processes, for example, such as shortening an application form, or changing the format of a survey, was highly unlikely to lead to relational learning. Many experiments were low-risk because there was no incentive for teams to take bigger ones. Learning "what works" therefore very much depended on the question being asked.

This is not to say that teams did not have bold ideas – some did. But the realization of imaginative experiments was only possible if they fell into the parameters of PCH's experimentation: only certain types of experiments were possible. The deeper learning associated with the normative and relational varieties was at odds with the tracked, measured nature of experimentation, and highlights challenges of learning when there are rules to be obeyed (Dunlop and Radaelli 2018).

Another factor that led to the low level of normative and relational learning was that experimentation prioritized certain types of evidence and results. Several teams spoke about the limiting nature of this:

> "… the type of change that you're trying to effect can't be counted immediately" (Interview N)
>
> "For me it's important to balance it by giving more importance to something that's qualitative, […] instead of focusing on 'we need a number'" (Interview A2)

While gathering insight in the form of comparable results and assessment of causality worked for some, for many teams it was a source of frustration.

A second finding with regards to the nature of learning was that learning did not translate to learning on a department-wide level, or a "shared learning agenda" (Sanders et al. 2020: 1042857). Communities of practice *did* develop. For example, in Experiment 20, an anticipated case of relational learning, the program team created a national working group around some of the themes of the experiment in order to share knowledge and best practice with partners on a national scale. Experiments 12 and 13, meanwhile, both still ongoing, created dialogue with sectoral partners in heritage institutions on issues related to education in the field. PCH's

Innovation and Experiment team also held open-invite sessions to gather feedback on and discuss experimentation. However, in general, engagement with communities of practice occurred only among those that were the most committed to (and, often, excited by) experimentation, and it was often informal:

> "At some point it's important for people to communicate lessons learned more largely in the department. […] you can't learn a lot if you don't share it with people" (Interview A2).

While recognizing the heterogeneous nature of the department's programs (and thus experiments), the danger is that that the learning from these experiments stays very localized to program teams and does not translate into a department-wide learning exercise, where there is more transformative potential for new strategies and policies to occur (Heikkila and Gerlak 2013).

The final dimension of learning to comment on is learning *how* to experiment. One interesting finding was that some individuals discussed learning a new skill set – how to approach problems differently. This was not the case for everyone and was dependent on the individual's background and experience. In a setting where employees are mobile and move between departments, there is the potential for longer-lasting capacity-building in experimentation methods, both within PCH and in the government more widely, if and when PCH employees move and 'take the learning with them.' However, this is ultimately dependent on leadership and political priorities; it does not matter if people know how to experiment if this approach is not supported by senior management and government.

6.2 What factors led to learning?

Beyond the nature of the individual experiment, there were several important contextual factors that influenced whether or not learning took place. The first was that working-level civil servants were given permission – by the highest civil servant in the department – to experiment and 'fail.' This was crucial; rational actors will not experiment without permission, because they will get in trouble if it goes wrong (see Howlett 2012): "[p]ublic servants are seeking clear signals that it's okay to take risks" (Privy Council Office 2019b: 15). Without this directive to experiment from the former Deputy Minister, such a robust program of experimentation would not have been created in PCH.

The second factor was that learning tended to be higher in program teams where there was a higher motivation for learning (Suškevičs et al. 2019). Experimentation was mandatory, but some teams were more eager and used it as an opportunity to do something ambitious and imaginative. As indicated above, the primary role of any program team is to *run the program*. There is little or no incentive to do things differently, especially for programs that have been running for decades with minimal changes. Teams that were more enthusiastic tended to be those that relished an opportunity to learn or change something; they were most positive about their experiences of experimentation and often ran higher-risk experiments and developed skills and confidence (Heiskanen et al. 2017).

Thirdly, experiments with stakeholder buy-in tended to lead to deeper learning. Experiments that were created with stakeholders, or involved stakeholders meaningfully, tended to lead to more deliberation and discussion, and therefore increased the potential for augmented trust and understanding between the department and stakeholders. This was the case with Experiment 9, in

which the experiment was created and administered with partners from another federal department and industry representations. In experiments 12 and 13, which were still ongoing, the team worked continuously with heritage institution partners and described the learning that occurred from these discussions.

The final success factor relates to the nature of the individual experiments. Experiments had to be appropriate to the political context – experimentation as defined by the department's and Government of Canada's parameters. A motivation to experiment was not enough: experimentation was not a *carte blanche* to do something different, or enact change; it was controlled and regulated. Proposals that did not conform to the rules set by the department and/or Treasury Board did not go ahead, and some teams had to modify their ideas in order to conform. The experiment therefore had to be appropriate, which meant not too risky, feasible, and with measurable results.

6.3 What lessons can be learned from this case study?

For learning to occur as a result of experimentation, it needs to be well-supported, properly resourced, and institutionalized in a way that it is not viewed as something transient or a pet project associated with one particular individual, a common perception in PCH. It also needed to have stronger overall commitment from senior management, *not* just the Deputy Minister. In order to have longer-term benefits – and particularly for multiple iterations of learning to occur (Sanders et al. 2020) – programs needed to have reassurance that the drive to 'do different' was not a short-term pursuit. The temporary nature of financing meant that teams struggled with how experimentation fit into longer-term program work. This under-resourcing therefore suggests a

disconnect between the transformative potential of experimentation espoused in policy documents, and the on-the-ground reality: many in PCH viewed experimentation as a *task to complete*. The agility and innovation expected in the top-down rollout of experimentation is unsustainable unless it has more bottom-up support in the form of buy-in and a sense of ownership among program teams.

Secondly, training in experimentation methods should come *before* the directive to experiment was given to program teams, which would have given teams the tools to identify problems rather than start with an instruction to experiment. Those looking to the future of experimentation in PCH at the time of interviews indicated that it will be more focused on starting with the identification of problems, "not a sweeping sort of top-down approach" (Interview Z). This is particularly important in order to build longer-term capacity in innovative methods. Moreover, evidence from interviews indicated that more support was necessary for corporate services. I also found many examples of technological challenges where inaccessibility to technology stalled some experiments or made others not feasible.

Thirdly, there was a lack of flexibility regarding experiment-related reporting requirements. There is a fundamental tension at play here: the innovative, less hierarchical thinking required of experimentation was constrained by the demands of the same approach – namely a "strengthened culture" of measurement and evaluation (Trudeau 2015). The centrally-driven and -coordinated innovation agenda meant an additional layer of reporting (Birch and Jacob 2019), which sat at odds with the reflection and deliberation necessary in experimentation. This is partially why there were so few examples of normative and relational learning: the 'shadow of hierarchy'

(Dunlop and Radaelli 2018) loomed large and many teams chose low-risk experiments. More time and space needed to be given in order to properly reflect on problems and putative outcomes (Farrelly and Brown 2011). Many people reported that they simply did not have the time to devote to experimentation. Learning from experiments requires reflection, which requires time.

Finally, there was also scope to improve communication and dialogue, both internally (with other program teams) and externally (to the stakeholders and general public). As indicated above, learning often did not spread beyond program teams. As Dunlop and Radaelli (2018: 260) argue, "dialogue doesn't just 'happen' and 'function' under any condition." There was potential to develop a stronger system of collective learning across the department, but this was not a core aspect of experimentation in PCH, and it only happened among the most engaged teams. Externally, considering the importance of the relationship with stakeholders, an overall communication strategy on experimentation may have led to fewer headaches for individual program teams. Future experimentation should also aim to involve stakeholders in as much of the experimentation process as possible, as early evidence suggests that this led to more instances of normative and relational learning through the building of stronger relationships. In addition, more communication with the general public about the benefits of experimentation can lead to increased visibility of and trust in the department's work (Rocle and Salles 2018).

In the end, structural factors both encouraged and detracted from learning. This case study showcases the challenges of trying to learn in a context that is dominated by hierarchical norms (Dunlop and Radaelli 2018) and a strong culture of ministerial responsibility and parliamentary

oversight. Overall, there is mixed evidence on whether the mandatory nature of experimentation was beneficial. While there was certainly resistance to experimentation among some programs and individuals, it is almost certainly true that optional experimentation would have led to much less learning. Were teams given the choice, however, this may have led to more normative and relational learning among teams that were particularly dedicated to it, if they were allowed more time and space for reflection and deliberation – particularly on identifying problems to tackle.

7 CONCLUSIONS

There has been long-standing interest in experimentation in policy-making, where its benefits have included observation of outcomes, learning, and flexibility (Dewey 1991[1927]). However, much less attention has been paid to the *processes* of doing experimentation and particularly on how the governance of experimentation can facilitate learning. The objective of this article was to examine policy experimentation in the context of policy learning in the Department of Canadian Heritage. In doing so it in part evaluated the Government of Canada's own goals in implementing an agenda of experimentation: to *learn what works,* in order to improve policy and program outcomes (Privy Council Office 2016).

Overall, the findings demonstrated that while there are barriers to implementing and operationalizing experimentation, it did lead to all three types of learning in the typology by Huitema et al. (2010). The most common type of learning was cognitive, new knowledge, but a small amount of both normative and relational learning also occurred. While the learning that took place went beyond the acquisition of new information, to (in a couple of cases) learning with more transformative potential, this was underdeveloped to date in this case.

The main barriers to learning stemmed from the hierarchical nature of how experimentation was governed (Dunlop and Radaelli 2018). Time and space for deeper reflection and exploration, and more secure resources, are crucial if the thinking encouraged during experimentation is to develop long-term. The dual aims of achieving innovation and delivering results, "while ensuring appropriate stewardship and accountability of public funds" (PCH 2018c: n.p.) work against each other: most program teams approached experimentation as a task to complete. This was further compounded by the additional challenge in PCH of programs' close connections with the sectors they fund. There is a fundamental tension between encouraging small-scale risk-taking and a broader institutional setting that is dominated by hierarchical norms. Ultimately, "structural impediments to change in many areas of the public service persists" (Privy Council Office 2019a: 2).

As many experiments were still ongoing when interviews took place, future research should study experimentation more comprehensively. In particular, the relationship between the type of experiment and learning could be studied in more detail. Other fruitful areas for future research are to ascertain whether any learning occurred among stakeholders outside of PCH and to compare PCH's experiences with other departments. The department's Report on Experimentation argues that the logic of experimenting and thinking differently "is becoming increasingly prevalent" (PCH 2018b: 21) in other areas of the department's activities such as policy development and performance management. However, again, political dynamics are crucial here. Prime Minister Trudeau was elected again in October 2019, albeit with a minority government. Experimentation is not mentioned in the 2019 ministerial mandate letters, and there

is a sense that the focus on deliverology and results-based outcomes has decreased (Brock 2021).

In addition, Ian Shugart, the new clerk of the Privy Council Office appointed in April 2019, has

indicated that he prefers cautious and responsible innovation; that "failure is something that

should be avoided" (Hoytema 2019). As such, the long-term implications of the Government of

Canada's possibly brief foray into policy experimentation remain to be seen.

SOURCES

Ansell, C. & Bartenberger, M. (2016). Varieties of experimentalism. *Ecological Economics,* 130(October), 64–73.

Ansell, C. & Bartenberger, M. (2017). The diversity of experimentation in the experimenting society. In I. van de Poel, L. Asveld, & D.C. Mehos (Eds.), *New Perspectives on Technology in Society: Experimentation Beyond the Laboratory* (pp. 36-58). London: Routledge.

Aucoin, P., Smith, J., & Dinsdale, G. (2004). *Responsible Government: Clarifying Essentials, Dispelling Myths and Exploring Change*. Ottawa: Canadian Centre for Management Development.

Baird, J., Plummer, R., Huag, C., & Huitema, D. (2014). Learning effects of interactive decision-making processes for climate change adaptation. *Global Environmental Change*, 27(July), 51-63.

Barber, M. (2015). *How to Run A Government: So that Citizens Benefit and Taxpayers Don't Go Crazy*. London: Penguin.

Birch, L. & Jacob, S. (2019). "Deliverology" and Evaluation: A Tale of Two Worlds. *Canadian Journal of Program Evaluation / La Revue canadienne d'évaluation de programme* 34(2), 303–328.

Bos, J.J., Brown, R.R., & Farrelly, M.A. (2013). A design framework for creating social learning situations. *Global Environmental Change*, 23(2), 398-412.

Brock, K.L. (2021). Policy labs, partners and policy effectiveness in Canada. *Policy Design and Practice*. Early view. Retrieved from https://www.tandfonline.com/doi/full/10.1080/25741292.2021.1880063

Brodkin, E.L., & Kaufman, A. (2000). Policy Experiments and Poverty Politics. *Social Service Review,* 74(4), 507-532.

Brown, H.S., & Cohen, M.J. (2019). Climate-governance entrepreneurship, higher-order learning, and sustainable consumption: the case of the state of Oregon, United States. *Climate Policy*, 19(6), 739-755.

Campbell, D. (1969). Reforms as experiments. *American Psychologist*, 24(4), 409-429.

Canada, Government of (2020) 2019 Public Service Employee Survey Results for Canadian Heritage. Retrieved from https://www.tbs-sct.gc.ca/pses-saff/2019/results-resultats/bq-pq/20/org-eng.aspx

Dewey, J. (1991[1927]). *The Public and Its Problems*. Ohio: Swallow Press.

Dunlop, C. & Radaelli, C. (2018). The lessons of policy learning: types, triggers, hindrances and pathologies. *Policy & Politics*, 46(2), 255-272.

Experimentation Works (2019, July 12) Experimentation Works Impact Report (2018–2019). Retrieved from https://medium.com/@exp_works/experimentation-works-impact-report-2018-2019-18434dd3d8fc

Farrelly, M., & Brown, R. (2011). Rethinking urban water management: Experimentation as a way forward? *Global Environmental Change*, 21(2), 721-732.

Heikkila, T. & Gerlak, A. (2013). Building a Conceptual Approach to Collective Learning: Lessons for Public Policy Scholars. *Policy Studies Journal*, 41(3), 484-512.

Heilmann, S. (2008). Policy Experimentation in China's Economic Rise. *Studies in Comparative International Development*, 43(1), 1–26.

Heiskanen, E., Hyvönen, K. Laakso, S., Laitila, P., Matschoss, K. & Mikkonen, I. (2017). Adoption and Use of Low-Carbon Technologies: Lessons from 100 Finnish Pilot Studies, Field Experiments and Demonstrations. *Sustainability*, 9(5), 1-20.

Hildén M., Jordan, A., & Huitema, D. (2017). The search for climate change and sustainability solutions - The promise and the pitfalls of experimentation. *Journal of Cleaner Production*, 169, 1-7.

Howlett, M. (2012). The lessons of failure: learning and blame avoidance in public policy-making. *International Political Science Review*, *33*(5), 539–555.

Hoytema, J. (2019, May 23). Privy Council clerk Ian Shugart urges public service to innovate — cautiously. *Ottawa Citizen*. Retrieved from https://ottawacitizen.com/news/local-news/privy-council-clerk-ian-shugart-urges-public-service-to-innovate-cautiously

Huitema, D., Jordan, A., Munaretto, S., & Hilden, M. (2018). Policy experimentation: core concepts, political dynamics, governance and impacts. *Policy Sciences*, 51(2), 143-159.

Huitema, D., Cornelisse, C. & Ottow, B. (2010). Is the jury still out? Toward greater insight in policy learning in participatory decision processes – the case of Dutch citizens' juries on water management in the Rhine Basin. *Ecology and Society*, 15(1), 16.

Mahler, J. (1997). Influences of Organizational Culture on Learning in Public Agencies. *Journal of Public Administration Research and Theory*, 7(4), 519–540.

McFadgen, B. (2019) Connecting policy change, experimentation, and entrepreneurs: advancing conceptual and empirical insights. *Ecology and Society,* 24(1), 30.

McFadgen, B. & Huitema, D. (2017). Stimulating Learning through Policy Experimentation: A Multi-Case Analysis of How Design Influences Policy Learning Outcomes in Experiments for Climate Adaptation. *Water,* 9(9), 1-22.

Mulcahy, K. (2006). Cultural policy: Definitions and theoretical approaches. *The Journal of Arts Management, Law, and Society*, 35(4), 319-330.

Newig, J., Kochskämper, E., Challies, E. & Jager, N.W. (2016). Exploring governance learning: How policymakers draw on evidence, experience and intuition in designing participatory flood risk planning. *Environmental Science & Policy*, 55(2): 353-360.

O'Brien, D. (2010). *Measuring the value of culture: A report to the Department for Culture, Media and Sport.*

Paquette, J., Beauregard, D., & Gunter, C. (2017). Settler colonialism and cultural policy: the colonial foundations and refoundations of Canadian cultural policy. *International Journal of Cultural Policy*, 23(3), 269-284.

Patrimoine Canada/Department of Canadian Heritage (PCH) (2020) Raison d'être, mandate and role. Retrieved from https://www.canada.ca/en/canadian-heritage/corporate/mandate.html

Patrimoine Canada/Department of Canadian Heritage (PCH) (2018a) 2018-2019 Departmental Plan. Ottawa: Canadian Heritage/Patrimoine canadien.

Patrimoine Canada/Department of Canadian Heritage (PCH) (2018b) A Plan for Experimentation at Canadian Heritage. Ottawa: Canadian Heritage/Patrimoine canadien. Retrieved from http://publications.gc.ca/collections/collection_2018/pch/CH1-40-2018-eng.pdf

Patrimoine Canada/Department of Canadian Heritage (PCH) (2018c). Departmental Evaluation Plan 2018-2019 to 2022-2023. Retrieved from https://www.canada.ca/en/canadian-heritage/corporate/publications/plans-reports/departmental-evaluationplan-2018-19-2022-23.html

Privy Council Office (2016). Experimentation direction for deputy heads. Retrieved from https://www.canada.ca/en/innovation-hub/services/reports-resources/experimentation-direction-deputy-heads.html

Privy Council Office (2019a) *Deputy Ministers Task Force on Public Sector Innovation: Report to the Clerk of the Privy Council 2018-2019*. Retrieved from https://www.canada.ca/content/dam/ih-ci/documents/pdfs/DM-TF-PSI-Report-AUG08.pdf

Privy Council Office (2019b) 2018-2019 Impact & Innovation Unit Annual Report. Retrieved from https://www.canada.ca/content/dam/ih-ci/documents/pdfs/rpt/impact-and-Innovation-Unit-2018-19-annual-report.pdf

Radaelli, C. (2003). The Open Method of Coordination: A New Governance Architecture for the European Union? Swedish Institute for European Policy Studies (SIEPS): Stockholm.

Rocle, N., & Salles, D. (2018). "Pioneers but not guinea pigs": experimenting with climate change adaptation in French coastal areas. *Policy Sciences,* 51(2), 231–247.

Sabel, C. & Zeitlin, J. (Eds.) (2010). *Experimentalist governance in the European Union: towards a new architecture*. Oxford: Oxford University Press.

Sanders, A.J.P, Ford, R.M., Keenan, R.J., & Larson, A.M. (2020). Learning through practice? Learning from the REDD+ demonstration project, Kalimantan Forests and Climate Partnership (KFCP) in Indonesia. *Land Use Policy* 91, 104285.

Stoker, G. (2010). Translating Experiments into Policy. *The ANNALS of the American Academy of Political and Social Science*, 628(1), 47–58.

Suškevičs, M., Hahn, T., & Rodela, R. (2019). Process and Contextual Factors Supporting Action-Oriented Learning: A Thematic Synthesis of Empirical Literature in Natural Resource Management. *Society & Natural Resources*, 32(7), 731-750.

Treasury Board of Canada (2016). Management Accountability Framework. Retrieved from https://www.canada.ca/en/treasury-board-secretariat/services/management-accountability-framework.html

Treasury Board of Canada Secretariat (2019). Population of the federal public service by department. Retrieved from https://www.canada.ca/en/treasury-board-secretariat/services/innovation/human-resources-statistics/population-federal-public-service-department.html

Trudeau, J. (2015, November 12). Mandate letter: Treasury Board of Canada. Retrieved from https://pm.gc.ca/eng/president-treasury-board-canada-mandate-letter-november-12-2015

Van der Heijden, J. (2014). Experimentation in policy design: insights from the building sector. *Policy Sciences*, 47(3), 249-266.

van Doren, D., Driessen, P.P.J., Runhaar, H.A.C., & Giezen, M. (2020). Learning within local government to promote the scaling-up of low-carbon initiatives: A case study in the City of Copenhagen. *Energy Policy*, 136, 111030.

Witting, A. (2017). Ruling out learning and change? Lessons from urban flood mitigation. *Policy and Society,* 36(2), 251-269.

Table 1: Typology of learning resulting from experimentations. Adapted from Huitema et al (2010), Baird et al. (2014), and McFadgen and Huitema (2017).

Type of learning	Definition
Cognitive	An acquisition of new knowledge
Normative	A change in perspectives, norms, and/or values
Relational	Increased understanding of others' mindsets, building of relationships

Table 2: Connecting experiments and learning

Experiment	Type of experiment	Type of learning
1	Administrative	Cognitive
2	Access	Normative
3	Administrative	Still in progress
4	Capacity	Still in progress
5	Access	Not enough evidence
6	Administrative	Cognitive
7	Access	Did not go ahead
8	Capacity	Cognitive
9	Capacity	Relational
10	Capacity	Cognitive
11	Administrative	Cognitive
12	Access	Still in progress
13	Access	Still in progress
14	Administrative	Cognitive
15	Administrative	Cognitive
16	Administrative	Cognitive
17	Administrative	Did not go ahead
18	Capacity	Did not go ahead
19	Capacity	Cognitive
20	Access	Still in progress
21	Access	Still in progress

22	Administrative / access	Still in progress
23	Capacity	Still in progress