

Translating Evaluation Policy Into Practice in Government Organizations

Isabelle Bourgeois  and Stéphanie Maltais 

American Journal of Evaluation
2023, Vol. 44(3) 353-373

© The Author(s) 2022



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/10982140221079837

journals.sagepub.com/home/aje



Abstract

Organizational evaluation policies describe how evaluation practices should be structured and implemented. As such, they provide key insights into organizational priorities and values regarding evaluation. However, the link between evaluation policies and how evaluation policies translate into concrete practices has seldom been explored until now. Our study examines the implementation of two Canadian federal government evaluation policies over a 10-year timespan, through the secondary analysis of reports produced on behalf of governmental evaluation functions. Our findings show that some policy elements have been fully implemented, but the way in which these have been implemented varies between organizations. Further, we observed that the level of control of various organizational members responsible for implementing policy elements, as well as time, can influence implementation of certain policy requirements. We conclude by proposing further directions for research to examine the policy-practice link.

Keywords

neutral assessment, evaluation function, Canada, evaluation policy, evaluation practice

Governmental evaluation functions are responsible for the development and implementation of studies to assess the relevance, effectiveness, and efficiency of policies and programs (House, Haug, & Norris, 1996). In many countries, a decentralized approach to evaluation has been implemented over time; this approach involves setting up evaluation functions that bring together evaluation experts responsible for coordinating external evaluations or conducting them internally on behalf of their organization (Lahey, Elliott, & Heath, 2018; Lemire, et al. 2018; Lemire, Nielsen, & Christie, 2018). Evaluation functions are typically guided by departmental and government-wide evaluation policies and thus serve as important mechanisms through which government evaluation priorities are brought into practice.

University of Ottawa, Ottawa, Canada

Corresponding Author:

Stéphanie Maltais, University of Ottawa, 120 University Private, Social Sciences Building, Room 8005, Ottawa, Ontario, K1N 6N5, Canada.

Email: smalt006@uottawa.ca

Despite their importance in directing and guiding the work of evaluation functions within governments, evaluation policies have seldom been examined empirically (Al Hudib & Cousins, 2022; Christie & Fierro, 2012; Christie & Lemire, 2019; Grob, 2006; House et al., 1996). Most notably, a need exists for knowledge and insight about how governmental evaluation policies are implemented by evaluation functions: "...limited empirical work has been conducted to better understand how they are interpreted and implemented by the evaluators and practitioners whose work they are likely to affect" (Christie & Fierro, 2012, p. 65).

Our study seeks to address this issue by examining the implementation of two Canadian federal government evaluation policies over the past 10 years, as seen through external assessments of departmental evaluation functions. More specifically, our primary objective was to determine the extent to which evaluation policy shapes departmental evaluation practice, as well as the facilitators and barriers to policy implementation in the departments. This work, although situated within a particular country's context, can inform the way in which government organizations around the world conceptualize and develop evaluation policies, and how these policies are then applied by departmental evaluation functions. Through our analysis, we will also probe the adequacy of existing typologies of evaluation policy and investigate specific areas of noncompliance or variations in policy implementation.

The paper begins with an overview of evaluation policy and its elements; we then provide a few examples of governmental evaluation policies. Next, we will describe the context for our study, including a detailed description of the two policies examined. Then, we will provide a description of our methods and present our findings. Finally, we will examine implications for policy development and implementation by evaluation functions.

Defining Evaluation Policy and its Elements

Mark, Cooksy, and Trochim (2009) define evaluation policy as a set of rules or principles embedded in legislation or other documents that are used to guide evaluation practice. These rules "help set the content, characteristics, and context of *evaluation itself*" (Mark et al., 2009, p. 5). Such policies can have varying degrees of formality and can be set through various means, depending on organizational context (Bicket, Hills, Wilkinson, & Penn, 2021; Mark et al., 2009; Trochim, 2009).

Evaluation policy can serve multiple purposes: first, it may guide the work of evaluators by identifying the types of programs and initiatives that must be evaluated, the frequency at which they should be evaluated, and the evaluation approaches and methods that should be used by evaluators (Christie & Lemire, 2019; Trochim, 2009). In this way, evaluation policy can guide as well as transform organizational evaluation practices by focusing on specific requirements and directing decision-making related to evaluation. Second, evaluation policy can be a communications tool, used to signal the importance of evaluation in supporting organizational learning, decision-making, and accountability (Trochim, 2009). Third, it may facilitate organizational conversations about evaluation by structuring expectations and requirements: "Establishing an evaluation culture requires work; it depends on a community of people who talk to each other about evaluation in a convivial and critical manner ... Expectations are important, especially that people should think about and do evaluation as a professional activity" (House et al., 1996, p. 139). This vision of policy reflects a top-down approach as conceptualized by Pressman and Wildavsky (1973) and requires organizational leaders and managers to implement policy requirements, thus translating the initial vision into concrete actions bound by resources and hierarchical control (Barrett, 2004; May & Wildavsky, 1978).

A taxonomy of evaluation policy areas proposed by Trochim (2009) identifies the various elements that can be included in evaluation policies. These refer to evaluation goals (purposes of evaluation), participation (stakeholder involvement in evaluation), management (how evaluations are managed within the organization), roles (in conducting and reviewing evaluations), process

(procedures required to conduct evaluation), methods (data collection and analysis), as well as evaluation utilization and capacity building. Along the same lines, Datta (2009), paraphrasing the 2007 AEA Policy Task Force, identifies seven areas of evaluation policy. These include: evaluation definition (distinguishing evaluation from other functions), requirements (what and when do we evaluate), methods (approaches for data collection), human resources (training, experience, and background of evaluators), budgets (financial resources required to evaluate), implementation (roles and responsibilities), and ethics. To these, she adds involvement and resource distribution toward evaluation capacity building. This latter point is also emphasized in the empirical work conducted by Al Hudib and Cousins (2022), which links evaluation policy to evaluation capacity building through several different connection points and which describes some of the moderating factors between the two. Dillman and Christie (2017) recognize the value in these typologies of evaluation policy but add a systems lens to broaden our understanding of how evaluation policies can guide evaluation practice. This view provides a better sense of the interconnections that can exist between various elements of the policy and how these interconnections can be leveraged to accomplish policy and organizational objectives.

Examples of Governmental Evaluation Policies

In the United States, government-wide evaluation policies include the Government Performance and Results Act (1993), which was revised and expanded in the form of the Government Performance and Results Modernization Act (2010). These policies emphasized performance measurement, as well as outcomes measurement, through evaluation activities (Lemire, Fierro, et al., 2018). More recently, the Foundations for Evidence-Based Policymaking Act (United States Congress, 2019) was signed into law in 2019 and seeks to improve the use of evidence and data in policymaking and program development across the federal government. Elsewhere, evaluation policies such as the United Kingdom's *Magenta Book* provide concrete guidance on evaluation practice, in order to enhance decision-making (Bicket et al., 2021). Others, such as the National Evaluation Policy Framework in South Africa, set the approach for establishing a national evaluation system on which provincial-level evaluations are conducted (Ishmail & Tully, 2020). All of these examples demonstrate governmental interest in evaluation policies and thus highlight the relevance of conducting research on their development and implementation.

In Canada as in the United States, the federal government is a key driver of demand for evaluation (Lahey et al., 2018; Lemire, Nielsen, et al., 2018). Requirements related to evaluation activities and reporting have been in place since 1977 in the Canadian federal government; over time, these requirements have evolved and have largely been subsumed within broader management and reporting trends and policies (e.g., New Public Management influenced the development of performance measurement and evaluation functions in the 1990s). Following two other evaluation-related policies, the *Policy on Evaluation* was implemented in 2009 across government, with some exemptions based on organizational size and mandate. This policy outlined the role of evaluation within departments and sought to improve the provision of “credible, timely, and neutral information on the ongoing relevance and performance of direct program spending” (Government of Canada, 2009, n.p.) to Ministers, central agencies, administrators, and the public. This policy was developed and implemented following a government-wide spending review in 2007, which sought to identify potential budgetary reallocations. The results of the review pointed to the fact that most departments and agencies did not have credible data on program relevance and cost-effectiveness, and that evaluation studies only covered a small percentage of programs. The Policy on Evaluation sought, among other things, to ensure that all direct program spending was evaluated every 5 years, in order to provide timely and accurate information on programs when needed by senior managers (Dumaine, 2012).

The Policy on Evaluation was rescinded in 2016 with the adoption of the *Policy on Results*. This broader policy replaced several management policies and sought to improve the achievement of results across government, following a change in the political environment. Evaluation remains a key component of the Policy on Results, but is linked to requirements related to performance measurement and program management. Notable changes include the requirement to evaluate all direct program spending, which was dropped, given its implementation challenges in the previous period. It also provides more flexibility in terms of issues that must be addressed as part of evaluations led by departmental evaluation functions, in order to support the production of relevant and timely evaluations that feed into the department's decision-making processes.

Linking Evaluation Policy to Practice Through “Neutral Assessments”

In 2009, the Policy on Evaluation required, for the first time, that large federal departments and agencies conduct a “neutral assessment”¹ of their evaluation function every 5 years (Government of Canada, 2009). This requirement was reiterated in the Policy on Results (Government of Canada, 2016) and, thus, has been in place for more than 10 years. The neutral assessment can be conducted in various ways and should focus on the compliance of the evaluation units with policy requirements. In other words, neutral assessments are meant to investigate how these two policies have been implemented by evaluation functions and how they have influenced evaluation practices within federal government organizations.

The neutral assessments provide a unique measure of the interplay between evaluation policy and practice; although a rich body of empirical research has been developed regarding the extent to which evaluations are used to account for public expenditures and to support organizational learning, systematic assessments of organizational evaluation functions are rare. In other words, “The *use* of evaluation has been studied empirically, but the *production* of evaluation has not” (House et al., 1996, p.136).

Our study, therefore, sought to capture the rich data described in neutral assessment reports to answer the following questions:

1. How have the elements of the Policy on Evaluation (2009) and the Policy on Results (2016) been implemented by federal evaluation functions?
2. To what extent do these policies appear to shape evaluation practice in the Canadian federal government? What are some of the barriers faced by evaluation functions in implementing evaluation policy?

The specific case of the Canadian federal government constitutes, in our view, a valuable contribution to the ongoing discussions about the policy-practice link occurring across governments and other types of organizations. To the best of our knowledge, the neutral assessment requirement in place in the government of Canada is unique and provides us with a distinct view of how evaluation policy is implemented in a real-world setting.

Conceptual Framework

The conceptual framework for our study is drawn from the typologies proposed by Trochim (2009) and paraphrased by Datta (2009), as well as from the Policy on Evaluation (2009) and the Policy on Results (2016). Our objective in designing this conceptual framework was to capture the general and specific elements that constitute evaluation policies, so that these can inform our analysis of the neutral assessment reports. In other words, our conceptual framework defines and describes the policy statements that are then implemented by evaluation functions. These elements are summarized in Table 1 and described in the section that follows.

Table 1. Conceptual Framework.

Trochim (2009)	Datta (2009)	Policy on Evaluation (2009) and Policy on Results (2016)
Evaluation goals	Evaluation definition	Evaluation coverage and issues
Evaluation participation	Evaluation involvement	Evaluation project management
Evaluation management	Evaluation budgets	Evaluation function management
Evaluation roles	Evaluation implementation	Oversight of evaluation function
Evaluation process	Evaluation ethics	Evaluation project management
Evaluation methods	Evaluation methods	Evaluation quality
		Integration of performance measurement
Evaluation utilization		Evaluation dissemination
Evaluation capacity building	Human resources Resource distribution	Evaluation utilization
		Competencies of heads of evaluation and evaluators

Evaluation coverage and issues: Refers to the issues and questions that frame evaluations. For example, policy statements in this area referred to relevance, effectiveness, and efficiency as primary evaluation issues.

Evaluation project management: Refers to the management of evaluation projects, led by senior evaluators (also called evaluation managers, team leaders, etc.).

Evaluation function management: Refers to the overall management of the evaluation function (also called directorate, division, branch, or unit), led by the Head of Evaluation.

Oversight of evaluation function: Refers to administrative structures set up to review evaluation plans and activities within the department or agency and advise the Deputy Head (most senior public servant) on evaluation-related activities; these oversight structures are typically called “Departmental Evaluation Committee” or “Performance Measurement and Evaluation Committee.” This element also refers to the roles and responsibilities of the Deputy Head in regard to evaluation.

Evaluation quality: Refers to the technical quality of the evaluation design, process, and report. Evaluations conducted in the federal government need to meet certain quality standards in order to support decision-making and accountability at the departmental level.

Integration of performance measurement: Refers to the availability, quality, utility, and use of performance information to support evaluations. This involves integrating program output and early outcome data, as well as other sources of administrative data, into evaluations to avoid duplication of efforts and to promote greater efficiency in the evaluation process.

Evaluation dissemination: This refers to the production and dissemination of evaluation reports to support evaluation use and public accountability.

Evaluation utilization: Refers to the use of evaluation reports and recommendations to improve existing programs and to support decision-making within the organization.

Competencies of heads of evaluation and evaluators: Refers to the professional competencies required by evaluators, as well as institutional supports provided, to develop these competencies.

The specific elements contained within each of these categories are introduced in the findings section and directed our data analysis by providing a coding structure for the content analysis. We were not able to directly associate some of the elements taken from the federal policies to the Trochim (2009) and Datta (2009) typologies. In a few instances, such as *Oversight of evaluation function*, we approximated these linkages. This suggests that further work on establishing a typology of evaluation policy elements might be needed, either to clarify some of the earlier work, or to add new dimensions to existing typologies.

Methods

Materials and Procedures

In collaboration with the Division on Results of the Treasury Board Secretariat, a call was issued to all Heads of Evaluation, asking them to provide neutral assessment reports published by their department or agency on a voluntary basis. In total, 41 reports were submitted from 28 different federal organizations. All of the reports were published between 2013 and 2020. The extent to which this sample of reports represents the entire population of neutral assessments produced over the study period is difficult to establish; changes in department names and functions are frequent in the Canadian federal government, and only certain departments and agencies are required to conduct neutral assessments, based on their size and mandate. In some cases, departments did not share their neutral assessment reports given concerns related to national security. All reports obtained were kept in a secure repository and a tracking document was created to capture data on year of publication and department or agency. This information is not shared here upon the request of some organizations that did not want to be identified as part of the study.

The way in which neutral assessments have been conducted over the past 10 years is an important issue that deserves further explanation, in order to judge the validity of the assessment reports as a data source. In 2014 and in 2018, the Treasury Board Secretariat issued guidance on how to conduct neutral assessments. According to the most recent version of the document, neutral assessments should demonstrate whether the evaluation unit under study complies to the various aspects of the Policy on Results (Government of Canada, 2016) and should offer actionable recommendations in the case of noncompliance. It does provide flexibility, however, in terms of the approach used to conduct the assessment, including: (a) an audit-like requirement-by-requirement assessment of compliance, (b) an assessment by theme that includes a review of governance, management practices, evaluation practices, and evaluation use, (c) a values-based assessment that could include credibility, independence, and the usefulness of evaluation products, or (d) a risk-based and targeted assessment that might focus on certain requirements, selected according to various risk factors. The guidance also provides flexibility in terms of the assessment period—in other words, neutral assessments can cover the entire 5-year period since the previous assessment or focus on a shorter period of time for some specific aspects of the policy. As in the case of program evaluations, neutral assessors can choose data collection methods, data sources, and analytical frameworks. Finally, by its very definition, the neutral assessment should be conducted by experts who do not have a direct relationship with the organization. Again, the guidance offers some flexibility in terms of who can conduct the assessment, and final approval of the report rests with the Deputy Head of the organization. An in-depth description of how the neutral assessments have been conducted, based on findings from this study, can be obtained in the study by Bourgeois and Maltais (2022).

Research Design

Our study used a nonexperimental, qualitative research design to extract secondary data from the 41 reports. The analysis was conducted in NVivo, according to the guidelines for *concept coding* qualitative analysis set out by Miles, Huberman, and Saldaña (2014). This specific approach involves using the exact words of participants (in our case, these were sentences or word groups extracted directly from the reports). Data extraction was guided by the conceptual framework presented earlier. The main categories correspond to the policy elements captured in the “Policy on Evaluation and Policy on Results” column in Table 1. More precise codes were developed based on the specific policy statements ($n = 44$). Each of these codes was defined and associated to examples pulled from the reports in order to create a codebook. In total, 1,278 data points were extracted

from the 41 reports by one team member, with validation checks conducted by the other team member.

Data Analysis

After the extraction was completed, we examined the data within each code for quality assurance purposes. Minor modifications were made after this examination and involved moving some data points between codes within the same category. A thematic analysis of the data contained within each code was then conducted in order to reduce and interpret the data. This analytical stage involved a careful review of the extracted data for each code, in order to highlight recurring issues or aspects. This work was conducted in accordance with best practices for using qualitative secondary data outlined by Ruggiano and Perry (2019). Our approach to data analysis supports a new use for the data contained within individual neutral assessment reports, which are usually not shared outside of the organization and not aggregated with corresponding data from other departments.

Analytical Quality and Methodological Limits

Various mechanisms were implemented to ensure methodological quality and validity (Creswell, 2014). For example, the team members communicated their understanding of the subject and engaged in peer debriefing to clarify data and analysis results at all stages of the study. This helped ensure clarity in the codebook. The main methodological limit of the study was, as described earlier, that we had an incomplete set of neutral assessments, since the Heads of Evaluation submitted their reports voluntarily. In addition, although we analyzed the methodology used to conduct each neutral assessment (Bourgeois & Maltais, 2022), we could not control or monitor the quality of the methodological approaches and findings presented in the neutral reports. However, we attempted to mitigate this limitation by extracting evidence and examples, rather than opinions, from the neutral assessment reports. This evidence does pose some additional limitations, since the original assessors chose the elements that they felt relevant in their reports and may have left other elements out of their analysis.

Findings

The overall picture obtained through our analysis is that, by and large, the two federal evaluation policies have been implemented by all of the organizations featured in the reports examined. The neutral assessments have served their purpose in determining which of the policy elements were implemented as intended, and in highlighting variabilities and nuances specific to each organization assessed. More importantly, our analysis shows that some policy elements have been implemented to a greater extent than others, and that the policy elements closer to the locus of control of senior executives tend to be more fully implemented than those further down the executive line. The paragraphs that follow explain these nuances in detail and provide examples from the neutral assessment reports.

Policy Implementation Status

The first observation stemming from our analysis is that none of the categories, based on the previous typologies and identified in our conceptual framework, has been fully implemented in the past 10 years. It is only at the code-level of our analysis that we can see policy elements that have been implemented and that are stable in all of the organizations reviewed. These policy elements tend to focus on specific actions, such as the creation of an oversight committee, developing and approving the

Departmental Evaluation Plan, designating a Head of Evaluation, and submitting evaluation reports to the Treasury Board Secretariat. For instance, as seen in Table 2, the category “Oversight of evaluation function” can be divided into seven policy elements. Some of these elements, such as “Creation of oversight committee” and “Approval of DEP,” have been fully implemented by all of the organizations studied. Other policy elements, such as “Ensuring access to information needed for evaluation and performance measurement,” have not yet been fully implemented. In this particular example, the policy elements that have been fully implemented are under the direct purview of the Deputy Heads, who are accountable for reporting on these to Treasury Board officials on a regular basis. It is not surprising, therefore, that these policy elements have been fully implemented.

Table 2 describes our main findings for each of the policy elements and their current implementation status. In some cases, data on several policy elements are reported together for the purposes of brevity and to show the interrelationships between these elements.

Variability in How Policies Are Implemented

Beyond reporting on the extent to which these policy elements were implemented, the neutral assessments also describe *how* they have been implemented. For example, all reports reviewed confirm that an **oversight committee** has indeed been created and provides advice to the Deputy Minister on matters related to evaluation and performance measurement, including the approval of evaluation reports and the Departmental Evaluation Plan. They also explain that senior departmental executives sit on this committee, and in some cases, external members are also recruited to provide specific expertise in evaluation or on key issues of interest to the organization. These points confirm that the policy element has been implemented and add a description of how it has been implemented. One states for instance: “A formal Terms of Reference (TOR) for the committee...sets out its mandate and roles and responsibilities, paralleling the requirements and expectations for such a committee of the ...2016 Policy.” In a few cases, the oversight committee is also responsible for the audit function; here, requirements regarding membership and processes appear to be more stringent. The size of the committee was also raised in a few instances where it was felt that having too many members had a negative impact on the committee’s ability to provide a true challenge function. This example shows that *policy implementation is not necessarily done in the same way* in all organizations. This is not a criticism; rather, an adaptation of policy elements is required in many cases, given the variability that exists in organization size, mandate, and purpose within the federal government.

The **creation of steering committees or advisory groups** made up of key program stakeholders is thought to contribute in a positive way to the evaluation process and provides another example of how a seemingly straightforward policy element can be implemented differently from one organization to the next. In some departments, stakeholders included in steering committees were working-level program staff who understand the challenges and complexities associated with the intervention. In other departments, steering committees were mainly composed of senior-level officials with a view of how the evaluation would be used as part of key decision-making processes. Some steering committees were only struck in particular contexts: “For more complex evaluations, a working group (advisory committee) is used to provide advice to the evaluators at key milestones during the evaluation.” In either case, the committees were found to provide important, contextual information as well as a critical eye on instrumentation and analysis. But the impact of each type of steering committee to the evaluation process (therefore, meeting policy objectives) can be different.

As a core component of evaluation practice, **methodology** features prominently in the neutral assessment reports. In this instance, variability in how the policy is implemented in specific

Table 2. Key Findings Regarding Implementation.

Policy elements	Implementation status	Implementation details
<i>Oversight of evaluation function</i>		
Creation of oversight committee	Full	All reports confirm that an oversight committee has been created and provides advice to the Deputy Minister (lead administrator) on matters related to evaluation and performance measurement, including the approval of evaluation reports and the Departmental Evaluation Plan.
Approval of departmental evaluation plan	Full	Early neutral assessment reports indicate that departments struggled with full coverage requirements; the Departmental Evaluation Plans identified a 5-year, full coverage objective, but in reality, few were able to fulfill this promise due to emerging priorities and resource allocation.
Coverage requirements	In progress	The Deputy Head is responsible for designating a Head of Evaluation and ensuring that this individual has unencumbered access and the freedom to raise emerging issues related to evaluation in the organization. This has been fully achieved as per policy requirements.
Designating head of evaluation	Full	This policy requirement seeks to ensure that the Head of Evaluation and the oversight committee have access to the information needed to assess departmental programs. In practice, this refers mainly to performance measurement data collected directly by the programs. The neutral assessment findings reveal that this continues to be a challenge for most departments, although there appears to be some progress on this front over the last 10 years.
Ensuring access to information needed for evaluation	In progress	In addition to approving evaluation reports, the oversight committee is also required to monitor the implementation of evaluation recommendations over time. Again, this seems to be improving, with more recent reports observing formal monitoring processes than older reports.
Monitoring and follow-up on evaluations	In progress	Evaluators tend to be regarded by their organizational peers as highly ethical. Their behavior tends to respect public service ethics and codes of conduct. By and large, evaluators do not find themselves in a position of conflict of interest very often and when they do, they are able to declare it through standard organizational processes.
<i>Competencies of evaluators</i>		
Public service ethics	Full	The competencies of evaluators were assessed positively in the reports reviewed. It should be noted, however, that most reports assessed evaluator competencies on the sole basis of stakeholder impressions. Only a few reports
Integrity	Full	
Conflicts of interest	Full	
Technical competencies	In progress—requires more complete assessment	

(continued)

Table 2. Continued.

Policy elements	Implementation status	Implementation details
Professional development	Full	presented information related to evaluator training and background as demonstrations of competency in evaluation. Some neutral assessments refer to a formal process linked to human resource management, such as an annual review and learning plan, while others identify previous training opportunities. These practices are reflective of the broader public service context and are not necessarily tied to evaluation policy.
Human resources	In progress	The neutral assessment reports commented extensively on the human resources dimension of evaluation function management. Most of the comments referred to a high turnover rate within the functions and the impacts of turnover on the function's ability to deliver on the Departmental Evaluation Plan.
<i>Evaluation function management</i> Sound fiscal decision-making	Not assessed	Only one report commented directly on sound financial management within the evaluation function; this could indicate that financial management was not of significant concern when the neutral assessments were undertaken, or that this is beyond the scope of such an assessment. <i>We flag this here to highlight that this may be one area where policy was not explicitly implemented through specific practices.</i>
Contracting	In progress	There exist differing views in terms of whether evaluations should be conducted internally or externally. These views are reflected in the neutral assessment reports, where some recognize the contributions of internal evaluators to contextually-appropriate evaluations that enrich organizational understanding and learning, and others highlight the importance of external views to achieve independence and accountability. More recent reports highlight the use of "hybrid teams," thought to provide the best value.
Conducting the departmental evaluation planning (DEP) process	Full	This policy requirement was completed by all departments included in the analysis. The annual DEP process is seen as an opportunity to consult across the department in order to include emerging issues and evaluation needs. It also provides the Head of Evaluation with supporting evidence to request additional resources for the function and to comment on past efforts.
<i>Evaluation project management</i> Timeliness of evaluation process	In progress	The timeliness of evaluations was a key consideration in all of the reports reviewed. Although some reports showed an improvement, others

(continued)

Table 2. Continued.

Policy elements	Implementation status	Implementation details
<p>Cost-effectiveness of evaluation process</p>	<p>In progress – needs further assessment</p>	<p>emphasized ongoing challenges that affect evaluators' ability to complete projects on time. The discussion of timeliness was often linked to evaluation scoping, which influences the overall project timelines depending on the nature of the evaluation questions posed and the data collection methods selected to answer them. Many reports examined the evaluation function's budget over a span of a few years, in order to determine whether significant increases or decreases could be attributable to the function's management processes. No clear trend emerged from the data extracted from the reports. Some reports focused on project cost-effectiveness, and identified various mechanisms that support the production of high-quality evaluations at a lower cost.</p>
<p>Peer review, advisory or steering committee groups</p>	<p>In progress</p>	<p>Program stakeholders involved in advisory or steering committees were found to enrich the evaluation by providing key pieces of information and reviewing deliverables at various phases of the process. These committees are not in place in all departments or for all projects.</p>
<p>Ad hoc consultation and expert input</p>	<p>Full</p>	<p>Most of the neutral assessment reports reviewed indicate that ad hoc consultations also occur and include program stakeholders as well as external subject matter experts.</p>
<p>Quality assurance</p>	<p>In progress</p>	<p>Quality assurance practices and procedures are variable. At a minimum, the Head of Evaluation reviews all evaluation reports. Some departments have taken this further and have developed quality control checklists and templates, and include a formal review process at various stages of the evaluation project.</p>
<p>Monitoring and follow-up by evaluators</p>	<p>Full</p>	<p>The reports tend to focus on monitoring related to the implementation of evaluation recommendations through the Management Response and Action Plan (MRAP) included in all evaluation reports. In a few cases, reports also refer to project post-mortems to inform and improve future evaluation projects.</p>
<p><i>Evaluation coverage and issues</i> Consideration of government-wide policies such as gender-based analysis and official languages</p>	<p>In progress</p>	<p>For the most part, the neutral assessment reports indicate that these issues are covered by evaluators, although they are not necessarily part of every evaluation project. The government's emphasis on GBA + in recent years has, however, been integrated to evaluation processes through the development of evaluation indicators, tools and training materials.</p>

(continued)

Table 2. Continued.

Policy elements	Implementation status	Implementation details
<i>Evaluation quality</i> Context, purpose, and objectives of evaluation	Full	<p>The neutral assessment reports do not raise significant issues about the context, purposes and objectives of the evaluations reviewed. For the most part, the evaluation reports include this information, and given the government of Canada's propensity for utilization-focused evaluation, the purpose and objectives of evaluations are typically to fulfill funding obligations or to inform senior executives on the relevance, efficiency and effectiveness of the program under study, as per policy requirements.</p> <p>Several neutral assessments examined project scoping approaches and results. Reports published after 2016 highlight the flexibility of departments in focusing on core issues within the more precise framing of the organization's decision-making process or to address program risk more efficiently.</p> <p>The reports indicate that evaluations are mainly founded on traditional methods, such as interviews, surveys, and document reviews, and some recommend that novel methods be considered to expand the function's repertoire of methodological tools. The neutral assessment reports also highlight the use of mixed methods and triangulation. Some of the reports are more critical and focus on methodological rigor and limitations. The most important observation regarding this policy requirement, however, is that none of the neutral assessment reports comment on evaluation approaches.</p>
Complexity and scope of evaluation	In progress	<p>The nature of the findings related to report parsimony tended to fall along two lines: the first indicated the high quality of the reports and the satisfaction of senior executives with the material; the second also highlighted the quality of the evaluation reports, but critiqued the length and technical nature of the reports. This poses a challenge to evaluators, who must produce a succinct report while also clearly linking evidence, findings, conclusions and recommendations.</p>
Evaluation approach and methods	In progress	<p>Findings and conclusions are perhaps the most important components of the evaluation reports. Many neutral assessment reports complimented the quality of the findings and conclusions sections of the report, often citing senior executives' opinions of these elements. Critiques pertained mostly to the substantiation of the findings with clear evidence, which was not always sufficient for the assessors.</p>
Parsimony	In progress	<p>Findings and conclusions are perhaps the most important components of the evaluation reports. Many neutral assessment reports complimented the quality of the findings and conclusions sections of the report, often citing senior executives' opinions of these elements. Critiques pertained mostly to the substantiation of the findings with clear evidence, which was not always sufficient for the assessors.</p>
Findings and conclusions	In progress	<p>Findings and conclusions are perhaps the most important components of the evaluation reports. Many neutral assessment reports complimented the quality of the findings and conclusions sections of the report, often citing senior executives' opinions of these elements. Critiques pertained mostly to the substantiation of the findings with clear evidence, which was not always sufficient for the assessors.</p>

(continued)

Table 2. Continued.

Policy elements	Implementation status	Implementation details
Recommendations	Full	<p>Evaluation recommendations are a cornerstone of Canadian federal evaluation practice. Recommendations are mandated by policy and must be included in all evaluation reports. For the most part, the neutral assessment reports state that the recommendations included in evaluation reports are clear and actionable. One area for improvement mentioned in a few cases is the need to better align recommendations with current program context and resources.</p>
<p><i>Integration of performance measurement</i> Implementation of performance measurement activities in organization</p>	In progress	<p>The integration of performance measurement in evaluation activities is an ongoing challenge for the function. Earlier neutral assessments (prior to 2016) often raise this issue, given that performance information was rarely available in the past and was deemed to be of low quality. The Policy on Results (2016) introduced specific requirements regarding performance measurement, including a change in oversight from a Departmental Evaluation Committee to a Performance Measurement and Evaluation Committee.</p>
<p><i>Evaluation dissemination</i> Submission of evaluation reports to Treasury Board Secretariat</p>	Full	<p>One of the main policy requirements regarding evaluation dissemination is that all approved evaluation reports must be submitted to the Treasury Board Secretariat within a specific timeframe of 120 days. Most of the neutral assessments revealed that this is done as requested and raise no significant issues. A few reports indicated that delays sometimes occur in the transmission process, although these do not appear to be of significant concern to the assessors.</p>
Public access to evaluation reports	In progress	<p>The 2009 Policy on Evaluation required that complete evaluation reports be published on the department's website no more than 120 days following approval by the Deputy Head; some neutral assessment reports indicate that delays often prevent this from happening within established timeframes. Since 2016, departments still publish their reports online, but are encouraged to also use other mechanisms, such as infographics, to disseminate their work publicly. Although many departments have done this, none of the neutral assessments explicitly presented information on alternative report formats.</p>
<p><i>Evaluation utilization</i> Internal decision purposes</p>	In progress	<p>At the program level, evaluations tend to elicit a positive response and a high rate of utilization. This was not always the case at a more strategic level. The</p>

(continued)

Table 2. Continued.

Policy elements	Implementation status	Implementation details
Treasury Board decision purposes	Full	<p>oversight committees tend to discuss evaluation findings within a broader organizational context, and in this way, evaluations can influence strategic decision-making; however, these conversations tend to be more general in nature and evaluations cannot be directly attributed to specific decisions. Evaluations were often cited as contributing to funding submissions to central agencies; this is done by design rather than happenstance, since the submission process typically requires evaluation data to support the request.</p>

evaluation projects is expected. The most important observation regarding this policy requirement, however, is that none of the neutral assessment reports comment on evaluation **approaches** such as utilization-focused evaluation or theory-based evaluation. This is somewhat surprising, given the publication of guidance on theory-driven evaluation in recent years (Government of Canada, 2021); we would have expected the neutral assessments to offer a more critical view of the choice and application of evaluation approaches. This is one example of partial implementation: evaluation functions respect basic principles in terms of methodology but tend to conflate approach and methodology.

Evaluation **recommendations** are a cornerstone of Canadian federal evaluation practice. Recommendations are mandated by policy and must be included in all evaluation reports. The neutral assessments confirm that recommendations are developed as intended; however, some identify a need to better align recommendations with current program context and resources. In some cases, recommendations were found to be unrealistic or unattainable. Managers felt, in those cases, that the evaluator's lack of understanding of the program context resulted in recommendations that could not be implemented as written. Therefore, the examination of neutral assessment reports reveals that although this particular policy element can be considered as fully implemented, the quality of implementation (i.e., actionable recommendations that take program context into account) can vary between organizations.

Evaluation **utilization** is another area in which policy has been partially implemented. Aside from accountability purposes, evaluations should also inform decision-making at the program level and more strategically at the organizational level. At the program level, evaluations tend to elicit a positive response and a high rate of utilization. However, the assessments were more critical of evaluation utilization beyond programs. The oversight committees tend to discuss evaluation findings within a broader organizational context, and in this way, evaluations can influence strategic decision-making; however, these conversations tend to be general in nature, and evaluations cannot be directly attributed to specific decisions. It should be noted, however, that most neutral assessment reports tackled evaluation utilization in a very broad way and did not identify specific examples of how evaluations led to concrete changes in a program or in an organization. Instead, they relied heavily on stakeholder interviews, through which they obtained general assurances of use rather than direct examples, as in the following passage: "Interviews indicated that the Deputy Minister and many Assistant Deputy Ministers (ADMs) have used evaluation information in program decision-making. Some ADMs could cite specific examples of their use of evaluation findings." A more rigorous methodological approach within the context of each neutral assessment might have provided clearer data on evaluation utilization at the organizational level.

Time and Level of Control Influence Policy Implementation

The preceding overview of findings highlights the time dimension in policy implementation: not all policy elements can be implemented at once, and some take longer than others to translate into concrete practice. Our findings show that the *level of control* of those tasked with implementing policy elements, among other factors, can influence what policy elements are implemented and when they are implemented. As mentioned previously, elements that fall directly under the purview of the Deputy Head tend to be implemented more quickly than other elements which may be under the responsibility of the Head of Evaluation. Another example pertains to the **timeliness of evaluations**, which was a key consideration in all of the reports reviewed. The neutral assessments determined that, although some departments showed improvement in terms of the time required to complete evaluation projects, others emphasized ongoing challenges that affect evaluators' ability to complete projects on time: "Foremost, four informants stated that, while the quality and quantity of information is there, the accessibility to timely information is not always ideal. Given the time lapse between the

collection and production of data—particularly for large, strategic evaluations that are carried out over a number of years, it can sometimes result in data and results that are obsolete.” In some cases, delays were caused by factors *external* to the function, such as contracting practices within the department, the approval process for evaluation reports, and the responsiveness of program staff and managers. In other cases, the evaluation function’s own *internal* practices were found to be the main cause of delays. For example, stakeholder involvement in evaluation was generally perceived as a positive factor, even though this created longer project timeframes and introduced some uncertainties in terms of project deliverables. The discussion of timeliness was also often linked to evaluation scoping, which is controlled for the most part by the evaluation function. Scoping influences the overall project timelines, depending on the nature of the evaluation questions posed and the data collection methods selected to answer them. To address this issue, several reports recommended more detailed planning and the search for greater efficiency throughout the evaluation project life-cycle. These examples illustrate how policy implementation depends on who is responsible for specific policy requirements and how much control they may have over the process.

Barriers Faced by Evaluation Functions in Implementing Policy

The neutral assessment reports highlight specific barriers in implementing evaluation policies. These barriers, like some of the other factors presented previously, often fall outside of the purview of the evaluation functions. One example of an enduring challenge faced by federal evaluators relates to the **integration and use of performance information** collected by programs or other organizational units. The policies state that high-quality performance measurement data should be accessible to evaluators in order to reduce the need for primary data collection and to capitalize on existing efforts and resources. Earlier neutral assessments (conducted prior to 2016) often highlight the difficulties faced by evaluators in accessing performance data, given its lack of availability and low quality. In response to these difficulties, the scope of the Policy on Results was broadened, introducing more stringent requirements related to reporting on the state of performance measurement in the organization. Neutral assessment reports published since then note some improvement but indicate that efforts must continue in order to improve the availability and quality of performance information for evaluation purposes: “The reality is that there continues to be a need to improve performance measurement. Indeed, it has been noted that ‘one-third to one-half of all recommendations in evaluation reports are around performance measurement...’”

Another barrier to policy implementation is the availability and retention of qualified evaluators, who are directly responsible for translating policy into practice. Many neutral assessment reports commented extensively on this issue by referring to a high **turnover** rate within the evaluation functions and its impact: “The increasing demand for evaluations, combined with a shortage of financial and human resources, has made it such that the Evaluation Division has faced challenges in meeting departmental needs that support effective and timely decision-making.” In some cases, turnover was addressed through external contracting, but this also caused significant challenges in terms of knowledge retention within the department.

How Practice Can Also Shape Policy

Although our findings generally show that Canadian federal evaluation policies have largely been translated into concrete practices by the evaluation functions, the reverse is also sometimes true: evaluation practices have influenced certain aspects of evaluation policies. The most striking example of this is related to “coverage requirements.” Coverage requirements refer to the ability of the evaluation function to evaluate (or “cover”) all required programs within a specific timeframe. These requirements have varied over the years; in the 2009 Policy on Evaluation, organizations were required

to evaluate 100% of direct program spending every 5 years.² The neutral assessment reports published between 2009 and 2016 indicate that departments struggled with these coverage requirements. Specifically, the Departmental Evaluation Plans identified a 5-year, full coverage objective, but in reality, few were able to fulfill this promise due to emerging priorities and available resources. In the 2016 Policy on Results, this requirement was changed to reflect organizational capacity and informational needs. In other words, the more recent policy backtracked the earlier requirement and recommended instead that evaluators focus on program risk and upcoming funding decisions to select the programs that should be evaluated. This particular issue provides an interesting example of how practice influences policy rather than the other way around.

Discussion

This descriptive study highlights the importance and role of evaluation policy in shaping evaluation practices in organizations, as defined by Christie and Lemire (2019) and Trochim (2009). The two evaluation policies examined are meant to guide the work of evaluators and the practice of evaluation at various levels of the organization. The policies do so by describing the role of senior executives in creating and supporting an evaluation function, the role of evaluation managers in developing an evaluation plan and ensuring that evaluations are conducted appropriately, and the role of evaluators in designing and conducting evaluations. In this way, these evaluation policies appear to support the establishment of an organizational evaluation culture as described by House et al. (1996).

Our findings confirm that the taxonomies of evaluation policy areas proposed by Trochim (2009) and Datta (2009) include elements found in other jurisdictions. Our analytical framework, based on the evaluation policies adopted since 2009 in the Canadian federal government, provides a deeper dive into these taxonomies by identifying specific policy elements that are one step closer to evaluation practice. These elements also support the argument put forth by Dillman and Christie (2017) about the need to describe the connections that exist between the elements in a systems-oriented analysis. We found, in some cases, that policy elements were difficult to distinguish when looking at practice (e.g., coverage of core evaluation issues is discussed at the organizational level and at the evaluation project level but refers to the same practice).

Our findings describe not only the extent to which policy elements have been translated into practice, but more importantly, how this was done in different organizations. They also bring to light the level of control and time required to implement certain elements and the barriers faced by evaluation functions during implementation. These findings are important in that they enable us to think of policy implementation not as one broad action, but as a series of specific acts dependent on various stakeholders and organizational context. They echo Lipsky's "Street-level Bureaucrats," meaning that managers and evaluators interpret and implement the policy elements differently based on their context and characteristics (Winter, 2012). Finally, our findings also show how practice can influence policy changes over time, and how this reverse policy-practice link is important to sustainable organizational improvement. This is consistent with the "bottom-up" approach described by Sabatier (1986).

The findings presented in this paper are unique in that they are founded on policies and practices implemented over a 10-year timespan (and in fact, may reflect more than 40 years of federal evaluation policy and practice), in organizations sharing a similar structure and with similar accountability requirements. This is a departure from previous efforts presented by Trochim (2009), Datta (2009), and Dillman and Christie (2017) but is close to the multiorganizational work conducted by Christie and Fierro (2012) and Christie and Lemire (2019). Despite their contextual framing, our findings also help illuminate the scalable link that can exist between evaluation policy and practice in organizations of all shapes and sizes. Our findings may also provide some food for thought to organizations interested in making this link more explicit in their own context.

Although this was not the original intention, our study also points to significant progress in several areas of evaluation practice since the implementation of the Policy on Evaluation in 2009. Many reports have applauded the attention given to enhancing staffing practices and the professional development of evaluators in recent years, as well as the efforts deployed by evaluation functions to support the integration of performance measurement across their organizations. Along the same lines, the neutral assessments have also described changes in how evaluations are reported to better meet the informational needs of senior managers through brief, targeted documents written with the evaluation user in mind. All of these findings demonstrate an evolution of the evaluation function over time. Many of these changes may be attributable to the adoption of the Policy on Results in 2016, as well as an increased understanding of the role that evaluation can play in government organizations. Some questions remain, however, about the added value of the evaluation function and how it can best support decision-making at the organizational level. This is a fundamental issue that future iterations of the policy should address explicitly. Additionally, given that our analysis did not include all neutral assessments produced since 2009, and that not all evaluation functions have undergone a neutral assessment, our findings must also be understood within the context of available neutral assessments rather than be generalized to all departments and agencies of the federal government.

In terms of future directions in this field, we argue that the call to action issued by Trochim in 2009 continues to ring true. The challenges and opportunities mentioned more than a decade ago continue to apply to today's evaluation landscape, and we support and reiterate the need for continued research on the topic of evaluation policy. Our study contributes to the field by analyzing neutral assessment reports, which roughly correspond to the "Evaluation Policy Audits" suggested by Trochim. It may be time to consider other activities suggested, such as evaluation policy analysis and evaluation policy clearinghouses and archives, to facilitate sharing of best practices and provide a knowledge base for future policy development and harmonization across organizations.

Conclusion

This study sought to examine how evaluation policy translates into organizational practices. The analysis of neutral assessment reports produced over the last decade enabled us to extract data pertaining to whether and how policy requirements are implemented, as well as the barriers that can sometimes be faced by evaluators charged with policy implementation. Overall, the two most recent evaluation-focused policies adopted by the Canadian federal government have been or are in progress of being implemented across the organizations reviewed. Our analysis also showed that in some instances, practice (or the inability to fulfill policy intentions) can also influence policy. Although our findings are situated within a particular context, we argue that the question of policy implementation through enacted evaluation practices is relevant to governments and organizations all over the world, and we reiterate the need for continued research and analysis on this topic.

In his description of what constitutes evaluation policy, Trochim (2009) establishes a clear distinction between what he calls "Substantive Policy," or the national policies meant to address specific social or economic problems, and "Evaluation Policy," which provides guidance to evaluators on the programs that should be evaluated and how evaluations should be conducted. Our findings confirm this role for evaluation policies, although we suggest that the body of knowledge on public policy and its implementation also applies to evaluation policy and should be examined further to identify lessons learned and best practices.

Finally, our work provides concrete examples of how evaluation policy is implemented in organizations, based on documentation developed for a different purpose. Although the neutral assessment reports were comprehensive in most cases, and focused on policy elements, we acknowledge

the limits of analyzing secondary data as part of our study. Future studies could delve deeper into the mechanisms of policy implementation in order to provide a more detailed picture of how policy is implemented into practice, as well as the role of policy in driving organizational evaluation efforts.

Acknowledgments

The authors wish to thank all of the Heads of Evaluation who provided their organization's neutral assessment reports, thus making this study possible. Their thanks also go to Cédric Ménard who provided useful background information and coordinated the request for the neutral assessment reports, and to Sebastian Lemire and the reviewers, who provided invaluable advice throughout the publication process.


Declaration of Conflicting Interests


The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Social Sciences and Humanities Research Council of Canada (grant number 435-2019-0725).

ORCID iDs

Isabelle Bourgeois  <https://orcid.org/0000-0002-2381-3093>

Stéphanie Maltais  <https://orcid.org/0000-0003-3752-0280>

Notes

1. The term “neutral assessment” is indicated in the policy and is the term in use in Canadian federal organizations to designate the policy implementation review exercise by a neutral party.
2. This requirement complements another legal requirement to evaluate all grants and contributions programs every 5 years, which was in place before the 2009 Policy on Evaluation was adopted.

References

- Al Hudib, H., & Cousins, J. B. (2022). Understanding evaluation policy and organizational capacity for evaluation: An interview study. *The American Journal of Evaluation, 43*(2), 234–254. <https://doi.org/10.1177/10982140211007573>
- Barrett, S. M. (2004). Implementation studies: Time for a revival? Personal reflections on 20 years of implementation studies. *Public Administration, 82*(2), 249–262. <https://doi.org/10.1111/j.0033-3298.2004.00393.x>. <https://login.proxy.bib.uottawa.ca/login?url=https://www-proquest-com.proxy.bib.uottawa.ca/scholarly-journals/implementation-studies-time-revival-personal/docview/60721402/se-2?accountid=14701>
- Bicket, M., Hills, D., Wilkinson, H., & Penn, A. (2021). Don't panic: bringing complexity thinking to UK government evaluation guidance. *Evaluation (London, England, 1995), 27*(1), 18–31. <https://doi.org/10.1177/1356389020980479>
- Bourgeois, I., & Maltais, S. (2022). Evaluating the evaluators: what have we learned from “neutral assessments” of the Canadian federal evaluation function? *Canadian Public Administration, 65*(1). <https://doi.org/10.1111/capa.12454>
- Christie, C. A., & Fierro, L. A. (2012). Evaluation policy to implementation: An evaluation of scientifically based research in practice. *Studies in Educational Evaluation, 38*(2), 65–72. <https://doi.org/10.1016/j.stueduc.2012.05.003>
- Christie, C. A., & Lemire, S. T. (2019). Why evaluation theory should be used to inform evaluation policy. *The American Journal of Evaluation, 40*(4), 490–508. <https://doi.org/10.1177/1098214018824045>

- Creswell, J. W. (2014). *Research design: qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.
- Datta, L. (2009). Golden is the sand: memory and hope in evaluation policy and evaluation practice. *New Directions for Evaluation*, 2009(123), 33–50. <https://doi.org/10.1002/ev.304>
- Dillman, L. M., & Christie, C. A. (2017). Evaluation policy in a nonprofit foundation: A case study exploration of the Robert Wood Johnson Foundation. *The American Journal of Evaluation*, 38(1), 60–79. <https://doi.org/10.1177/1098214016642864>
- Dumaine, F. (2012). When one must go: The Canadian experience with strategic review and judging program value. *New Directions for Evaluation*, 2012(133), 65–75. <https://doi.org/10.1002/ev.20007>
- Government of Canada. (2009). *Policy on evaluation*. <https://www.tbs-sct.gc.ca/pol/doceng.aspx?id=15024§ion=html>
- Government of Canada. (2016). *Policy on results*. Retrieved from <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=31300§ion=html>
- Government of Canada. (2021). *Theory-based approaches to evaluation: concepts and practices*. Retrieved from <https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/evaluation-government-canada/theory-based-approaches-evaluation-concepts-practices.html>
- Government Performance & Results Act of 1993, Public Law 103-62, 107 Stat.285.
- Government Performance & Results Modernization Act of 2010, Public Law 111-352, 124 Stat. 3866.
- Grob, G. F. (2006). The evaluator's role in policy development. *New Directions for Evaluation*, 2006(112), 99–108. <https://doi.org/10.1002/ev.211>
- House, E. R., Haug, C., & Norris, N. (1996). Producing evaluations in a large bureaucracy. *Evaluation*, 2(2), 135–150. <https://doi.org/10.1177/135638909600200202>
- Ishmail, Z., & Tully, V. L. (2020). An overview of the provincial evaluation system of the Western Cape Government of South Africa as a response to the evaluation of the National Evaluation System. *African Evaluation Journal*, 8(1), e1–e10. <https://doi.org/10.4102/aej.v8i1.425>
- Lahey, R., Elliott, C., & Heath, S. (2018). The evolving market for systematic evaluation in Canada. *New Directions for Evaluation*, 2018(160), 45–62. <https://doi.org/10.1002/ev.20346>
- Lemire, S., Fierro, L. A., Kinarsky, A. R., Fujita-Conrads, E., & Christie, C. A. (2018). The U.S. Federal evaluation market. *New Directions for Evaluation*, 2018(160), 63–80. <https://doi.org/10.1002/ev.20343>
- Lemire, S., Nielsen, S. B., & Christie, C. A. (2018). Toward understanding the evaluation market and its industry—advancing a research agenda: toward understanding the evaluation market and its industry. *New Directions for Evaluation*, 2018(160), 145–163. <https://doi.org/10.1002/ev.20339>
- Mark, M. M., Cooksy, L. J., & Trochim, W. M. K. (2009). Evaluation policy: an introduction and overview. *New Directions for Evaluation*, 2009(123), 3–11. <https://doi.org/10.1002/ev.302>
- May, J. V., & Wildavsky, A. B. (1978). *The policy cycle* Sage Publications. Retrieved from <https://login.proxy.bib.uottawa.ca/login?url=https://www-proquest-com.proxy.bib.uottawa.ca/books/policy-cycle/docview/38140956/se-2?accountid=14701>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications, Inc.
- Pressman, J., & Wildavsky, A. (1973). *Implementation: How great expectations in Washington are dashed in Oakland*. Retrieved from <https://login.proxy.bib.uottawa.ca/login?url=https://www-proquest-com.proxy.bib.uottawa.ca/books/implementation-how-great-expectations-washington/docview/60902928/se-2?accountid=14701>
- Ruggiano, N., & Perry, T. E. (2019). Conducting secondary analysis of qualitative data: Should we, can we, and how? *Qualitative Social Work: QSW: Research and Practice*, 18(1), 81–97. <https://doi.org/10.1177/1473325017700701>
- Sabatier, P. A. (1986). Top-down and bottom-up approaches to implementation research: a critical analysis and suggested synthesis. *Journal of Public Policy*, 6(1), 21–48. <https://doi.org/10.1017/S0143814X00003846>

- Trochim, W. M. K. (2009). Evaluation policy and evaluation practice. In W. M. K. Trochim, M. M. Mark, & L. J. Cooksy (Eds.), *Evaluation policy and evaluation practice. New directions for evaluation* (Vol. 123, pp. 13–32). San Francisco: Jossey-Bass.
- United States Congress. (2019). Foundations for Evidence-Based Policymaking Act of 2018. 115th Congress, Public Law 115–435—Jan. 14. <https://www.congress.gov/bill/115th-congress/house-bill/4174>
- Winter, S. C. (2012). Implementation perspectives: Status and reconsideration. In B. G. Peters & J. Pierre (Eds.), *The SAGE handbook of public administration* (pp. 265–278). London: Sage Publications Ltd. <https://www-doi-org.proxy.bib.uottawa.ca/10.4135/9781446200506.n17>