From New Public Management to New Political Governance: Implications for Evaluation

Jill Anne Chouinard  
Department of Educational Research Methodology, School of Education  
University of North Carolina at Greensboro

Peter Milley  
Faculty of Education  
University of Ottawa

Abstract: Public administration scholars have discerned a shift in the federal governance context in Canada, from what was traditionally a strong, nonpartisan public service to a more politicized, even partisan, model of public decision-making with power concentrated in the upper reaches of the political executive. We explore the potential implications of these changes for evaluation in the federal bureaucracy. Our analysis, tentative at this point, suggests that in light of heightened political pressures, and a decline in the use of “evidence” in federal policy circles, evaluations may present an increasingly complex activity for public administrators to manage. These developments raise important questions for the evaluation community about its relationships with public managers and its role and professional values in a democratic institution.

Keywords: evaluation use, new political governance, new public management

Résumé : Les observateurs du domaine de l’administration publique ont constaté un virage dans le contexte de la gouvernance fédérale au Canada, laquelle est passée d’une administration publique forte, non partisane, à un modèle plus politisé, voire partisan, de la prise de décisions où le pouvoir est concentré dans les plus hautes sphères de l’exécutif fédéral. Nous explorons les impacts possibles de ces changements sur l’évaluation au sein de la bureaucratie fédérale. Notre analyse préliminaire semble indiquer que, au vu des pressions politiques accrues et du déclin du recours aux « preuves » dans les milieux politiques fédéraux, l’évaluation devient une activité de plus en plus complexe à gérer pour les administrateurs fédéraux. Ces développements soulèvent des questions importantes au sein de la communauté de l’évaluation pour ce qui est de ses liens avec les gestionnaires du secteur public ainsi que de son rôle et de ses valeurs professionnelles au sein d’une institution démocratique.

Mots clés : utilisation de l’évaluation, nouvelle gouvernance politique, nouvelle gestion publique

Corresponding author: Jill Anne Chouinard, 252 School of Education, 1300 Spring Garden St., Greensboro, NC 27412; jill.chouinard@uncg.edu

© 2015 Canadian Journal of Program Evaluation / La Revue canadienne d’évaluation de programme  
30.1 (Spring / printemps), 1–22 doi: 10.3138/cjpe.30.1.1
INTRODUCTION

People know what they do; they frequently know why they do what they do; but what they don’t know is what they do does. (Foucault (1982) as cited in Dreyfus & Rabinow, 1982, p. 187)

As in much of the West, evaluation in Canada is a key part of the public governance architecture. Despite the lack of perceived benefits of evaluation in the 1960s and early 1970s (Dobell & Zussman, 1981), evaluation has become an institution in society (Dahler-Larsen, 2012), what others have termed an “explosion” (Power, 1997). Following the rise in the 1980s of New Public Management (NPM), evaluation has responded to the need for improved efficiency, effectiveness, and accountability, creating a vast technocratic network focused on meeting the demands of public sector managers (Norris & Kushner, 2007). The past 30 years have witnessed a substantial increase in evaluation, reporting, monitoring, inspection, and audit in all public sector organizations, led in part by pressure from international organizations such as the Organization for Economic Cooperation and Development and other international monitoring agencies (Leeuw & Furubo, 2008). In light of this “obsession” with evaluation and monitoring (Dahler-Larsen, 2012), the number of evaluation associations has grown from 1 in the 1980s (the Canadian Evaluation Association) to over 158 today. There has also been a sizeable increase in the number of evaluation journals, conferences, training opportunities, and consultancies. As Leeuw and Furubo (2008) keenly observe, evaluation systems spawn more evaluation systems.

The historical rise to prominence of evaluation in public governance in Canada, along with its future status, merit reflection in 2015, the International Year of Evaluation (as declared by the United Nations). Since the new millennium, profound changes have occurred in the governance context. Public administration scholars in Canada and other Anglo-American countries have been describing and debating some discernable shifts. At the federal level in Canada they have observed a concentration of power in the upper reaches of executive government (Savoie, 2008) and a more politicized tenor to the relationships between elected officials and federal public servants (Aucoin, 2012). This shift is seen to signal far-reaching negative implications for the workings of government (of which evaluation is a key part), including with respect to policy making and administration, transparency and public access to information, the production and use of research findings (or evidence), and democratic conventions and principles (Aucoin, Bakvis, & Jarvis, 2013; Heintzman, 2013, 2014; Kozolanka, 2014; Turner, 2013).

Inheriting an important raison d’être from the federal war on poverty initiated in the United States by Lyndon Johnson in the 1960s (Rossi, Lipsey, & Freeman, 2004), the North American evaluation community has long positioned itself in the role of promoting social betterment (Henry, 2000; Mark, Henry & Julnes, 2000) and providing relevant, up-to-date information to policy makers and to the public at large. In describing social betterment, Henry (2000) describes
Collectively we must decide what are the most pressing social problems, decide what actions should be taken to address them, and continue to refine the actions that are taken and reengineer the organizations that are charged with achieving these outcomes. (p. 90)

While evaluation maintains deep social, cultural, and historical roots, and remains fully integrated and embedded in the administrative architecture, the federal governance context in Canada is seen to be shifting to a more politicized, increasingly partisan (and perhaps less democratic) focus. Aucoin (2012) and others (Heintzman & Juillet, 2012; Jarvis & Bakvis, 2012) call this new dynamic the “New Political Governance” (NPG), observing that it diverges from NPM in significant ways. We argue that evaluation needs to be understood within this shifting matrix, one where political considerations infused with the partisan interests of the government of the day appear to be assuming greater prominence within the social and organizational context of the federal public service.

In this article we explore (from a conceptual point of view) the potential implications for evaluation of the governance trends prominent members of the public administration community have identified, as encapsulated in the NPG model. What influences or effects might these trends have on evaluation and on the function of evaluation within government? If evaluation is fundamentally about social betterment, what role might evaluation be expected to play in this new federal context? To explore these broad questions, the article is organized into four parts. First, we describe the changing context and approach in public governance as a shift from NPM to NPG. Next, we explore what the potential effects of this shift on evaluation might look like in terms of (a) the “hyper” politicization of the context for evaluation, (b) the use and influence of evaluation in the public service, (c) evaluation as an institutionalized practice in federal bureaucracies, and (d) the influence of evaluation in shaping public perception and public dialogue.

In the third section, we turn to a discussion of two possible paradoxes—or what Dahler-Larsen (2012) might term “constitutive effects”—of evaluation in this transitional moment. We conclude with a brief discussion of future implications for evaluation practice in public sector contexts.

To our knowledge, the specific terrain of NPG explored in this article is new to the field of evaluation. Media commentators have picked up on the trends encapsulated in NPG, applying them to a range of issues, such as the centralization of power (Jarvis & Turnbull, 2012), increasing partisanship (Simpson, 2014), and government communications and advertising (Mendes, 2014). We note that media coverage is scant in terms of evaluation and surmise this is because evaluations are largely performed at program levels and are less likely to be scrutinized as they are deemed less relevant to system-wide concerns (Shepherd, 2012). In terms of monitoring functions in the federal context, the parliamentary, media, and public focus has instead been primarily on “watchdogs” and “czars,” such as the Parliamentary Budget Officer, the Auditor General of Canada, and the Information Commissioner of Canada. Senior officials associated with these offices have been publicly undermined by the government of the day when they announced
findings that contradicted those of the government or provided commentary or information that allowed the government’s policies or actions to be seen in a negative light (Beeby, 2014; Harris, 2014). It has been argued that the symbolic messages from cases such as these have resulted in the growth of a “culture of fear” and “self-censorship” in the public service, undermining its conventional roles of providing impartial advice and speaking truth to power (Heintzman, 2014).

With respect to evaluation and NPG, we are thus on new territory. As a result, we approach this inquiry as a speculative exercise grounded in a spirit of exploration that draws inspiration from the concepts and arguments advanced in the evaluation literature by those offering sociological and critical perspectives (e.g., Dahler-Larsen, 2012; Power, 1997). What follows can, in some ways, be considered as a “foresight study” in that we are asking: What seems to be going on? What might happen? What might we need to do (Giaoutzi & Sapio, 2013)? The value in such an exercise “lies less in forecasting accuracy than in its usefulness … in opening minds to consider new possibilities, and thus change the policy agenda” (Giaoutzi & Sapio, 2013, p. 4). We thus advance three goals: (a) begin a dialogue on NPG between the public administration community and the evaluation community, (b) prompt reflection and debate in the evaluation community related to the federal context of evaluation, and (c) lay the groundwork for lines of inquiry to be pursued through empirical research in the future.

FROM NPM TO NPG: A CHANGING CONTEXT AND APPROACH IN PUBLIC GOVERNANCE

From the early 1980s to late 1990s, public administration scholars identified, analyzed, and debated a set of related reforms that came to be known as New Public Management (NPM) (Hood, 1991; Pollitt, 2013). The changes associated with NPM focused on improving responsiveness, efficiency, productivity, and service quality in the public sector (Aucoin, 2012). A concern about the behaviour of public administrators also lay behind this agenda, with some observers seeing public servants as more interested in securing and expanding mandates and budgets for their organizations than in serving the interests of citizens (Niskanen, 1991; Peters & Savoie, 2012). These administrative behaviours were seen to undermine both efficiency and democracy, generating two strategic countermoves on the part of elected officials in executive government. The first involved breaking the quasi-monopolistic roles public bureaucracies played in providing policy advice and delivering services. This was accomplished through such means as delayering departments, outsourcing delivery, using public-private partnerships, and encouraging networks for policy development and program coordination (Jarvis & Thomas, 2012). The second was to adopt and apply private sector management principles and practices (Aucoin, 1990, 1995; Hood, 1991; Pollitt & Bouckaert, 2004), including the extensive use of performance measurement and evaluation (Chouinard, 2013).

Looking back, three fundamental precepts existed at the heart of NPM (Aucoin, 2012). The first was to (re)locate the power for making public policy
decisions in the hands of democratically elected representatives, signalling the latter’s superiority in this area (Jarvis & Bakvis, 2012), including with respect to the management reform agenda (Peters & Savoie, 2012). To support this shift, prime ministers and presidents in Anglo-American jurisdictions increasingly concentrated power and resources in their offices and the central agencies of government—a particularly prominent development in Canada (Savoie, 2008).

The second tenet was to give front-line public managers and their agents more discretion in the delivery of programs and services. The theory was that this would allow them to address “client” needs better and to seek efficiencies at local levels. The third precept was to ensure that managers managed well and achieved desired results. This involved increased use of performance measurement, evaluation, auditing, and reporting (Peters & Savoie, 2012). The “insistence on results and evaluation” (Jarvis & Thomas, 2012, p. 293) privileged “accountability for control and assurance” over “accountability for learning” (Howard & Phillips, 2012, p. 322). More financial and process controls were also implemented along with more auditing (Hood, 2004; Power, 1997).

The Emergence of “New Political Governance” in the Canadian Context

Some aspects of NPM were not embraced as enthusiastically in Canada as they were in other jurisdictions (Jarvis & Thomas, 2012); however, the concentration of power and control in the office and person of the Prime Minister and central agencies of government has featured prominently in the Canadian context (Savoie, 1999, 2008). Public administration scholars argue five pressures in the governance context have motivated this change (Aucoin, 2012; Jarvis & Thomas, 2012). These pressures have elicited some identifiable responses on the part of political executives (see Table 1).

Table 1. NPG: Pressures in the Governance Context and the Response of Political Executives

<table>
<thead>
<tr>
<th>Pressures in the Governance Context</th>
<th>Response of Political Executives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ubiquitous media presence</td>
<td>Expansion of communications infrastructure and centralization of power and control</td>
</tr>
<tr>
<td>Increased demands for transparency, access, openness</td>
<td>Tactical manoeuvring to control information flows and access</td>
</tr>
<tr>
<td>Expansion of “watchdog” agencies and groups</td>
<td>Discrediting of challengers</td>
</tr>
<tr>
<td>Competition in policy marketplace</td>
<td>Enthusiastic promotion of chosen policies</td>
</tr>
<tr>
<td>Increased volatility in the electorate</td>
<td>Dualistic, partisan politics</td>
</tr>
</tbody>
</table>


doi: 10.3138/cjpe.30.1.1

CJPE 30.1, 1–22 © 2015
The first pressure point is the ubiquitous presence of an aggressive, negative media, including various publics using social media (Jarvis & Bakvis, 2012). In response, the federal government's communications infrastructure has burgeoned since the mid 2000s (Kozolanka, 2014). Some public servants working in this area have been placed in situations that undermine their actual or perceived impartiality (Aucoin, 2012; Kozolanka, 2014), while others, such as federal scientists, report being “muzzled” (Chung, 2013) and asked to “exclude or alter information” in their reports (May, 2014).

The second pressure is the call for increased transparency, openness, and public access to information. While largely seen as modern cornerstones of good governance, these developments have also created resource and political strains for governments (Peters & Savoie, 2012). Anecdotal evidence suggests some public managers have come under increasing pressure to commit less to writing and to restrict what is made public (Aucoin, 2012). Interference in the access to information process by ministerial staff has also been reported (Office of the Information Commissioner of Canada, 2014).

A third strain has been the significant expansion of “watchdog” agencies. In Canada, the public and media tend to respect and trust these agents, but actors in the political executive have sought to discredit them in light of partisan calculations (Simpson, 2014). Not surprisingly, the release of reports from these review agencies is tightly scripted on both sides.

A fourth development has been increased competition in the policy marketplace. The number of lobbyists and interest groups has grown and think tanks have emerged, some of which have a partisan bent but remain independent of official parties. Ministers and their staff call upon public servants to protect their interests in the face of these groups, which sometimes leads to unorthodox work in promoting and justifying government policy rather than discussing it in neutral and “purely” informative ways (Mulgan, 2007).

The fifth political stressor is increased volatility in the electorate, which is less deferential to elected officials, less loyal to parties, and less engaged. In response, political parties take polarized stances to win support and votes from “core” supporters. This dualistic politics affects public servants who, in serving one government, subsequently come to be seen as “suspect” by a different incoming government. Some appear to deal with this dilemma by exhibiting strong—or “promiscuously partisan” (Aucoin, 2012, p. 179)—support for the policies of each ruling party of the day, a democratically unconventional behaviour.

Aucoin (2012) coined the term New Political Governance (NPG) to describe changes he and others (Savoie, 2008; Heintzman & Juillet, 2012; Jarvis & Bakvis, 2012) discerned in political and administrative behaviours in light of these pressures. NPG reflects the ongoing concern among political leaders to reassert their “democratic right to govern by taking control of the state apparatus” (Aucoin, 2008, as cited in Jarvis & Bakvis, 2012, p. 16). However, the governance context has taken on such a political character that leaders are tempted “to do whatever they deem necessary to stay in power” (Aucoin, 2012, p. 181), including tightening
controls over the bureaucracy to increase “responsiveness and limit politically damaging errors” (Heintzman & Juillet, 2012, p. 343).

Four main features characterize NPG (see Table 2) as an “ideal type,” or heuristic model, against which the deviance or conformity of specific empirical cases can be assessed (Aucoin, 2012; Parsons, 1964; Weber, 1968). The correspondence of actual conduct to these features will depend on “the party in power, the prime minister, the state of competition between parties in the legislature and in the electorate, and, among other factors, the institutional and statutory constraints that provide checks against politicization” (Aucoin 2012, 179).

The first feature is the tendency for elected officials to engage in permanent electoral campaigning during their time in office, as reflected in the heightened desire of political leaders to centralize power and control over decisions and communications. If governments are highly partisan and politicized, there is a higher risk of misuse of this centralized power over the public service for partisan purposes (Aucoin, 2012). The second feature is the growing number and power of political advisors and ministerial staff. In their zeal to serve the ruling party, these political operatives can fall into the trap of regarding the conventional democratic values of a nonpartisan public service as “obstacles to overcome” (Aucoin, 2012, p. 186) and seek to co-opt public servants as political agents. The third feature is the temptation to politicize the senior ranks of the public service by appointing or promoting those who are seen to be “on board” with the agenda (Aucoin, 2012). In addition, even if appointments are not overtly politicized, with the centralization of power and control there is more pressure on senior public servants to put the interests of the Prime Minister and his or her staff above those of the public interest and their ministers. The fourth feature is that governments want public servants to be publicly supportive (i.e., “enthusiastic”) about policies and to promote (rather than describe or explain) them to citizens and stakeholders—even in parliamentary committees and hearings. The concern here is that governments are demanding

<table>
<thead>
<tr>
<th>Main Features of NPG</th>
<th>Related Risks to Democratic Conventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permanent electoral campaign</td>
<td>Misuse of centralized power over public service for partisan purposes</td>
</tr>
<tr>
<td>Growing number and power of political advisors and ministerial staff</td>
<td>Coercion or co-optation of public servants as political agents</td>
</tr>
<tr>
<td>Politicization of senior ranks of the public service</td>
<td>Undermining of principles of merit, professionalism, and neutrality in the public service</td>
</tr>
<tr>
<td>Expectation of “enthusiastic” response of public servants to (partisan) policy</td>
<td>Erosion of convention of impartiality in the public service</td>
</tr>
</tbody>
</table>

Table 2. Main Features of NPG and Related Risks to Democratic Conventions

partisan loyalty from public servants rather than impartial loyalty as has traditionally been the case.

Taken together, the pressures in the governance context and the NPG responses to them weaken the traditional norms of political neutrality, professionalism, and relative independence in the public service (Heintzman, 2013; Heintzman & Juillet, 2012). Jocelyne Bourgon (2014), former Clerk of the Privy Council, calls this an “institutional erosion” stemming in part from “a progressive de-professionalization … where a parallel political public service, comprised of people who are neither elected nor office-holders in the civil service, yield great power without accountability for the influence they exercise” (p. 8). Aucoin (2012) goes further in arguing NPG “constitutes a corrupt form of politicization to the extent that governments seek to use … the public service in the administration of public resources and the conduct of public business to better secure … partisan advantage” (p. 178), observing that this change in relationships and practices has the potential to erode public trust in government over the long term. The precepts and idealized relationships between ministers and public servants that underpin NPG build on those of NPM, but also diverge in significant ways as a result of increased power in the hands of the Prime Minister and her or his political staff and heightened emphasis on the political/partisan agenda.

In light of these changes, public administration scholars have mainly focused their attention on the effects of NPG on democratic conventions and accountability. A sustained critical gaze has not yet been cast on the potential implications of NPG for evaluation and on the role of evaluators in the federal context in Canada. Aucoin (2012) provides a starting point, however, in arguing the partisan politicization of the public service has negative effects on the performance of public managers and their organizations if, to assess NPG, one applies “nonpartisan criteria” (p. 178). This thought is suggestive of how evaluation can be subject to politicization—e.g., under NPG “high performers” are those who enthusiastically advance political objectives and outcomes while abandoning a commitment of speaking truth to power (Savoie, 2003). It is also indicative of how the “normal” political pressures that surround evaluation in government may be amplified and endowed with partisan hues under NPG, potentially generating similar issues regarding mandates, terms of reference, methods, access to information, communication and use of findings, independence, integrity, and professional values felt in other politically sensitive areas of government such as communications (Kozolanka, 2014), science (Chung, 2013; Franks, 2010; May, 2014; Turner, 2013) and review agencies (Beeby, 2014; Chase, 2013; Simpson, 2014).

FROM NPM TO NPG: POTENTIAL IMPLICATIONS FOR EVALUATION AND FOR THE WORK OF EVALUATORS

The point that we are noting is the discernable shift from NPM to NPG, from what was once a strongly democratic and nonpartisan public service to a more politicized and partisan model of public decision-making, may ultimately have
serious implications for the planning, implementation, and reception of evaluation in government. In this section we discuss four such possible implications of this shift: (a) the “hyper” politicization of the context for evaluation, (b) evaluation as institutionalized and formalized practice within federal bureaucracies, (c) the use and influence of evaluation in public sector contexts, and (d) the influence of evaluation in shaping public perception and public dialogue.

The “hyper” politicization of the context of evaluation. Evaluation and politics have always been intimately connected (Weiss, 1973), with evaluation considered the “handmaiden to political leadership” (Cronbach and Associates, 1980, p. 30). The increased emphasis on partisan political considerations in the federal context in Canada, along with the concentration of power at the centre of government, amplifies and alters the relationship of evaluation and politics. Under NPM, the connection between evaluation and public management was formalized in the Treasury Board of Canada’s results-based agenda and enshrined in policies such as the Management Accountability Framework (Treasury Board Secretariat, 1997) and the Federal Accountability Act (Treasury Board Secretariat, 2006), both of which helped establish evaluation as a key tool of public policy providing information on program performance and spending. The results-based agenda emphasized the instrumental use-value of evaluation in improving program management, planning, and decision-making, and in providing a means of control (Everitt, 1996). This agenda also played up the symbolic use-value of evaluation (Chouinard, 2013), as its widespread adoption signalled assurance to the public that the government’s political and organizational commitments were fiscally prudent.

In the “hyper” political context associated with NPG, evaluation may take on even more symbolic importance as a source of political legitimation (Taylor & Balloch, 2005) and, even more importantly, potential delegitimation. In an environment of permanent electoral campaigning, “spin” (MacDougall, 2015), and strict central control over policy and communications, political sensitivities are heightened throughout the public bureaucracy about how information generated about government policies and operations may be taken up in Parliament, the media, or the public sphere to criticize the government. Numerous cases have already come to light about how the production of and access to information (Kozolanka, 2014) and “evidence” (Harris, 2014; Turner, 2013) have been tightly circumscribed in the federal government since 2006, with some commentators noting this has led to a climate of fear (Chase, 2013; Chung, 2014). A prominent example has been the so-called “muzzling” of scientists who work for the federal government, which led to a “Death of Evidence” march on Parliament Hill in July of 2012 (Turner, 2013), and to 800 scientists from 32 countries publicly calling upon the government to end its “restrictions on scientific communication and collaboration faced by Canadian government scientists” (Chung, 2014). Another high-profile example is how the parliamentary budget officer, appointed in 2008 under the Accountability Act to provide nonpartisan analysis of the nation’s finances and government spending (Harris, 2014), struggled unsuccessfully from 2012 to 2013 to have the government release details on billions of dollars of budget cuts under the Deficit
Reduction Action Plan, eventually resorting to a federal court case (Chase, 2013; Heintzman, 2014). Although these prominent cases did not specifically concern evaluation, the political and symbolic dynamics associated with them may extend to evaluation processes as well. Of course, this general dynamic has always been part of the situation for evaluation, but in an NPG context we are concerned that it may be amplified dramatically and altered such that evaluators may find the “politics of information and evidence” are more focused on issues regarding the circumscription of mandate, scope, criteria, informants, access, and reporting than they are about methodological selection (e.g., quantitative versus qualitative data). The instrumental and symbolic use-value public managers placed on having scientifically rigorous evaluations under NPM may, under NPG, give way to concerns about how to address potential and actual political issues associated with the evaluation. The perceived demise of evidence-based practice under an NPG regime, coupled with more centralized political and bureaucratic decision-making, a public service that has been labelled “promiscuously partisan,” and more tightly scripted public messaging, may substantially reduce the instrumental use-value of evaluation, as its symbolic dimension as a source of political legitimation or potential delegitimation takes precedence.

Evaluation as institutionalized and formalized practice. The formal adoption and use of evaluation technologies within federal governance contexts, identified as mechanisms to assist decision-making, monitoring, reporting, improvement, learning, and control of programs and expenditures (etc.), is prevalent. Whether in education, healthcare, social services, or the financial sector, evaluation plays a legitimizing function within government. In fact, establishing the practice (and reputation) of evaluation is considered a key component of a well-functioning public service. As such, evaluation has become a key mechanism of the accountability movement (Chaytor, MacDonald, & Melvin, 2002; Norris & Kushner, 2007; Schwandt, 2009), with evaluators acting as key agents of public accountability (Greene, 1999; Stake, 2001).

Despite its technological and rational/logical façade and its apolitical cast, the current evaluation and performance management architecture within NPM and NPG is not neutral, as its representation extends beyond providing a mere technical solution to the current problems of public governance (Behn, 1998). As Power (1997) has posited, “audit is not simply a solution to a technical problem; it also makes possible ways of redesigning the practice of government” (p. 11). As we have argued previously (Chouinard, 2013), the current federal architecture, intended to guarantee compliance with the standards of efficiency and effectiveness, imposes central control, as government departments become predominantly focused on completing detailed performance measurement and evaluation plans to satisfy federal accountability requirements. In effect, these “technologies of control” have created an internal system of self-monitoring and self-surveillance, creating what Franklin (1990) referred to as a “culture of compliance,” and what Foucault (1972, 1977) might have identified as a form of disciplinary power. As a key part of our organizational culture, evaluations (and evaluative thinking) can thus be said to
structure what we talk about, how we frame what we determine as important questions and problems, and how we define possible solutions to these problems. As Dahler-Larsen (2012) notes, as organizations assume cultures of evaluation, our thinking thus becomes more organizational, essentially more bureaucratized.

This notion of central control extends beyond the sphere of technical activities and processes associated with evaluation to include the professional norms and standards that frame the parameters of evaluation practice. As the Canadian federal evaluation and accountability policies make clear, the purpose of evaluation is to generate “accurate, objective and evidence-based information to help managers make sound, more effective decisions on their policies, projects and initiatives” (Treasury Board Secretariat, 2001, Preface). The knowledge generated through evaluation thus tends to be framed in narrow, technical, and instrumental terms, and valued for its objectivity, predictability, and measurability. Under the current system of governance, knowledge claims that are associated with positivist/postpositivist methodologies are given greater credence (Henry, 2009; Steiner, Wroblewski, & Cook, 2009), helping circumscribe (and essentially define) what is considered “legitimate” (or official) knowledge. As Norris and Kushner (2007) so aptly surmise, “knowledge control lies at the heart of all New Public Management projects” (p. 3).

The changing Canadian context, from a focus on building a solid architecture to support the policies and practices of accountability (e.g., effectiveness, efficiency, and transparency) under NPM, to a more centralized and political platform for decision-making under NPG, may thus have profound implications for evaluation and for the federal bureaucracy. Under NPM, evaluation and performance management are seen as key features of the governance strategy, regardless of whether findings are used for strategic-level decisions (Shepherd, 2012). While still a key part of the federal architecture, evaluation under NPG seems to be shifting to a more symbolic role institutionally, shifting to what Dahler-Larsen (2012) has identified as a “functional alternative to action” (p. 87). The governance shift encapsulated in NPG may thus create a more tenuous link between evaluation and government actions and decisions. The more partisan and autocratic the political leadership is in an NPG-like context, the more tenuous that link may likely become.

The use and influence of evaluation in public sector contexts. Despite the fact that utilization is one of the most studied areas in evaluation (Christie, 2007), ever-changing and shifting political and economic contexts and ongoing challenges related to instrumental and direct use of findings ensure its endurance as a relevant area of concern. The notion of utilization remains central to evaluation (Vedung, 1999). The evaluation literature identifies a range and taxonomy of use (e.g., direct or instrumental use, conceptual use or enlightenment, political or symbolic use; see Patton, 1988; Weiss, 1998), with an emphasis on how evaluation can contribute to learning, capacity building, organizational development, empowerment, process use, and so on. Despite the myriad conceptualizations of use currently discussed and debated in the field, the value of evaluation nonetheless continues to be related to the expectation that evaluation results will in some way
directly inform the policy and the decision-making process (Weiss, 1999). Paradoxically, the increase in the demand for evaluation does not seem to have led to increased utilization, leading Dahler-Larsen (2012) to identify this phenomenon as one of the “mysteries” of the evaluation wave. This “mystery” is not relegated to the use of evaluation findings alone, as under NPG we also note the disuse or misuse of science and social science (so-called “evidence”) within government as well (Prewitt et al., 2012; Turner, 2013). Given the politicization of decision-making within the federal government under NPG, evaluation use (and the use of social science research more generally), is becoming far more complex. As Weiss (1999) has cautioned, the more democratic the system, the more likely that evaluation results will be used, while authoritarian systems tend to “hold power close to the chest and fewer people are allowed into its inner sanctum” (p. 480).

While we previously noted the role of evaluation within the current federal system as largely symbolic (Chouinard, 2013), the institutionalization of evaluation within the federal bureaucracy is ensuring that its impact will extend beyond the symbolic, to influence how we work and how we think about our work (Dahler-Larsen, 2012). In many ways, the utilization issue has moved beyond whether evaluation results are being used, to encompass the more subtle (and yet more far-reaching) impact of the institutionalization of evaluation within federal bureaucracies. Although evaluation has always been understood as a cultural, social, and political expression embedded in and intrinsically tied to particular social and institutional structures and practices (House & Howe, 2000), the shifting context from NPM to NPG seems to emphasize for us the fundamental transformation in conceptualizations of evaluation use.

If, under NPM, evaluation has played an important symbolic role with respect to accountability, over time it has also changed what public managers work on and how they perceive and go about their work. The institutionalization of evaluation suggests progress has been made in promulgating evaluative thinking among public managers, a longstanding goal of the evaluation community. But, in a more profound sense, it signals how public managers have been encouraged to internalize the discipline of evaluation (e.g., planning, monitoring, and course correction in light of external direction and standards) as part of their legitimate participation in the broader “culture of compliance” in the public service. The arrival of NPG may further entrench this dynamic and amplify its political and partisan dimensions, as public managers likely know “full well” that the central agencies and the Prime Minister’s Office have them on “constant watch to protect the government’s political interests” while having the authority and resources “to intervene to set things right” (Peters & Savoie, 2012, p. 34). This suggests the goal the evaluation community has pursued to push public managers to learn conceptually may be overtaken under NPG by a steep learning curve with respect to politics. This may be particularly true for senior civil servants who “are now doing many things that political assistants used to do in the past” (Bourgault, 2013, p. 171).

Under NPG, the symbolic dimension and use of evaluation may well be more pronounced. On a landscape where the promise for the use of findings seems
dim because the desire for “immediate political advantage” casts “long shadows over hard data and reasoned analysis” (Turner, 2013, p. 17), evaluation processes and results may increasingly be subject to tight scripting by overinvolved public managers in service to the senior ranks of the public service. The politicization of the process may ensure that we move to a focus on the performative nature of public management rather than a focus on government performance, a shift in the metaphor for evaluation from a technical, methodological focus to representing the imagery of performance and dramaturgy.

The influence of evaluation in shaping public perception and public dialogue. Thinking of evaluation as a form of “assisted sense making” (Mark, Henry, & Julnes, 2000) draws attention to the fact that evaluation is not merely a technical activity, but one that has the potential to enhance our understanding of a program, process, policy, or practice. The sense-making concept also contributes to our aspirations for evaluation, to its role and function in democratic communities linked to “social betterment,” to the utility and use of evaluation, and to the shaping of our concepts and measures as they relate to the parameters of program understanding. In many ways, evaluation has become a fundamental part of our public policy model and culture (Dahler-Larsen, 2012), as it helps frame our understanding of public programs in terms of how they are defined, how they work, and how they can be made more effective and efficient. The current focus of evaluation on measures of efficiency, effectiveness, and value for money (as defined within a Canadian context) shapes our understanding of programs and possibilities, ultimately shaping the parameters of public expectation, dialogue, and discourse in ways that give the illusion of democratic accountability (Schwandt, 2009).

Historically there has always been a strong link between evaluation and the machinery of government, with evaluation serving an instrumental role designed to assist with managerial decision-making and program improvement (House, 1993). Under NPM, and as the current Canadian public governance model well illustrates, evaluation has evolved to an accountability or largely symbolic function used to legitimize government activities (Chouinard, 2013; Norris & Kushner, 2007; Schwandt, 2009). The noted shift to NPG, highlighted by tight control of information and communication, along with an increase in the number of unelected political staff within public bureaucracy, may mean that public dialogue and discussion around public programs and government spending priorities will remain tightly scripted and highly controlled.

POTENTIAL PARADOXES (OR “CONSTITUTIVE EFFECTS”) PRODUCED WITHIN SHIFT FROM NPM TO NPG: IMPLICATIONS FOR FURTHER DISCUSSION

This article highlights the complexity of social, political, and institutional forces involved in knowledge and policy formation, as well as the role of diverse actors and systems implicated in producing different types of knowledge for the
management of the public good. As Weiss (1999) makes so abundantly clear, the creation and construction of policy is never straightforward, particularly given that the policy-making arena is characterized by an array of different and often competing agendas and political interests about what is required and how best to organize and allocate existing resources. Evaluation findings (and data from the social sciences more generally) provide merely one construction amidst a cacophony of other competing interests and alternative constructions available. Understanding the relationship between evaluation, utilization, and public policy is therefore complex and, as our analysis has thus far highlighted, contains numerous contradictory or paradoxical dynamics and tensions. In using NPG as a heuristic model with which to explore contemporary dynamics in public governance, we identify two potential paradoxes with respect to the work of evaluators and with respect to the function of evaluation more generally: the paradox of use and the paradox of accountability.

We use the term *paradox*, which the *Oxford Dictionary* identifies as a self-contradictory proposition and a senseless or logically unacceptable conclusion, as a way to problematize the push-pull dynamics and tensions that may exist for evaluators and for evaluation under NPG-like conditions. Our use of the term paradox is intended to illustrate the embeddedness of these tensions or contradictions within the very processes and outcomes of public governance, what Dahler-Larsen (2012) might refer to as “constitutive effects.” The important point to keep in mind is that constitutive effects describe the effect of evaluation systems on constructing a particular social reality (e.g., while we use tests to measure student learning, the testing itself changes what is taught, how it is taught, and ultimately what students learn). The following discussion is thus intended to describe the potential paradoxes and constitutive effects of evaluation during the key transitional moment from NPM to NPG.

The paradox of use refers to the fact that, despite the discourse surrounding evidence-based practice and evidence-based decision-making, the actual use of evidence seems to be in decline, while the infrastructure surrounding evaluation (what we have referred to as the evaluation industry) seems to be on the rise. In fact, evaluation is woven into the very fabric of public governance and concomitantly forms an obligatory component of virtually all aspects of policy, service delivery, and management processes—from micro to macro levels. Yet, arguably, evaluation use within public sector discourse and decision-making seems to be on a downward trend. If this is true, what are the implications of this paradox for evaluation and for evaluators?

From our perspective, the most profound effect of the paradox of use might be the potential influence on the learning and pedagogical function of evaluation. Although evaluation has traditionally been conceptualized as a pedagogical undertaking (Schwandt, 2003), under NPG its use may well be reduced to a performative function (almost a political dramaturgy) that fundamentally changes not only what we learn but also how we learn it. The distinction here is between evaluation as a “technical undertaking” intended as a set of tools designed to
improve practice and generate answers, and evaluation as a conceptual practice intended to generate dialogue and facilitate learning (Schwandt, 2003). Learning would then be focused on tracking and measuring performance and progress against stated goals, rather than as an educative instrument for capacity building and dialogue. Over time, the paradox of use likely shifts to a “performance paradox,” as the connection between performance indicators and performance itself becomes weaker (van Thiel & Leeuw, 2002).

The paradox of accountability refers to the inclusion of evaluation as a core component of the democratic and management accountability discourses (Chaytor et al., 2002; Schwandt, 2009), where it is seen to provide “assurance” about results (primarily) and processes (secondarily), while in practice it seems to remain largely unclear who is accountable to whom and for what. Although few would dispute the importance of accountability in public governance, what it means in theory and practice remains vague. Charlton (2002) describes accountability as a “slippery rhetorical term” (p. 3) containing two distinct, interchangeable meanings: one related to technical management processes and the other to democratic accountability. Aucoin and Jarvis (2005) argue it is the latter form of accountability that seems particularly elusive, despite being a mainspring of good governance. Heintzman & Juillet (2012) point out tensions that exist with respect to democratic accountability at the interface of the political executive and public administration, all of which may become more problematic with the emergence of NPG. Of particular concern is a weakening of the “key tenets of traditional regimes, such as the political neutrality, professionalism, and relative independence of the public service” (Heintzman & Juillet, 2012, p. 344).

As discussed earlier, evaluation under NPM was positioned as a technical instrument positioned to provide assurance that public managers managed well and to provide a means for controlling and altering their behaviours. It was also seen to support democratic accountability by helping elected officials to wrest control for policy and decisions away from bureaucrats and providing evidence to Parliament and the public about the effectiveness of policy and its implementation. These aspirational goals were not fully realized in practice, but they nonetheless helped evaluation rise to institutional prominence.

Our analysis thus far suggests that under NPG the symbolic use of evaluation may come to the fore, and the control and disciplining function it provides in the public administration will have a more pronounced political (and potentially partisan) flavour. Evaluation may present a more complex problem for public administrators to manage under NPG than it did under NPM, as it would seem to be much more than a technical process aimed at generating information about management performance; rather, it may become a political concern that cuts straight to the heart of democratic conventions. This potential shift raises important questions for the evaluation community about its relationships with public managers and its role and professional values in a democratic institution.
CONCLUDING NOTES AND AGENDA FOR FUTURE RESEARCH

To this day, the boom in evaluation continues, even as its ability to realize its ambitions (providing evidence to inform decisions, supporting learning processes, bolstering democratic and management accountability, helping improve efficiency and effectiveness) seems remote. In this article we highlighted the shift from a public sector governance model based on the principles of NPM to those of NPG, with a focus on the potential implications of this shift for evaluation (and the role of evaluation) within a changing federal bureaucracy. While evaluation continues to be a key part of public governance, the politicization of federal policy and decision-making, coupled with a decline in the use of "evidence" in federal policy circles, suggests the performative, dramaturgical aspect of evaluation may take precedence over the more direct use of findings. The institutionalization of evaluation within the federal bureaucracy also highlights its disciplinary, self-monitoring function, as federal decisions are increasingly made beyond the range of the federal bureaucracy, and ultimately beyond the conventional parameters of our democratic structures.

In this article we argued that evaluation, as a form of assisted sense-making and fundamentally related to raising issues and questions of social betterment, is now at risk. In an important sense, we may be witnessing a shift in governments’ perspectives on the main purposes of evaluation along MacDonald’s (1974) typology, from aspirations for democratic aims of “social betterment” in the 1970s, to bureaucratic utility and functions of government under NPM in the 1980s and 1990s, to the emergence of autocratic expressions in the 2000s, as signalled by a more politicized, potentially less neutral or nonpartisan, public service. As evaluation seems to be increasingly positioned as part of the tightly scripted political drama (a “reality series”?) that unfolds inside and outside the public service, we have to ask: If evaluation is fundamentally about social betterment, what role can it be expected to play within this new federal context? What role can evaluators working within public sector institutions be expected to play? Perhaps being aware of the shifting context within the federal bureaucracy may enable evaluators to become more politically savvy in their engagement in a process that has become so much more politicized, a process whose connections to social betterment we argue may well be at risk.

As we noted, this article is intended as a “foresight study,” as a way to generate dialogue among the evaluation community (and perhaps between the evaluation community and the public administration community), and as a speculative piece to prompt reflection on the shifting context for evaluation within the federal bureaucracy. In our minds, this article lays the conceptual groundwork for future empirical research that we propose to explore the potential effects of the shifting context (from NPM to NPG) on the function and use of evaluation within the federal bureaucracy. What discernable changes can be identified to describe the shifting context for evaluation? How is evaluation currently perceived, and how is it being used? What is the relationship between evaluation and departmental decision-making? Our goal in pursuing empirical research at this time is to
identify current evaluation practices and emergent patterns of use within what we have noted is a shifting, more politicized federal context.

REFERENCES


Giaoutzi, M., & Sapio, B. (2013). In search of foresight methodologies: Riddle or necessity? In M. Giaoutzi & B. Sapio (Eds.), *Recent developments in foresight methodologies* (pp. 3–9). Boston, MA: Springer. http://dx.doi.org/10.1007/978-1-4614-5215-7_1


doi: 10.3138/cjpe.30.1.1

CJPE 30.1, 1–22 © 2015


**AUTHOR INFORMATION**

Jill Anne Chouinard is Assistant Professor in the Department of Educational Research Methodology, School of Education, University of North Carolina at Greensboro. She received her PhD in organizational studies and evaluation from the University of Ottawa in 2015.
2010. Her main interests are in cross-cultural/culturally response evaluation, qualitative research, and participatory research and evaluation. Given her interest in using evaluation as a leverage for social change, her research and writing also focuses on the relationship between evaluation and public policy. She has extensive experience working on evaluations at the community level in the areas of education and training, social services, health, and organizational learning and change.

Peter Milley is Assistant Professor of Leadership, Evaluation, Curriculum, and Policy at the Faculty of Education, University of Ottawa. He completed his PhD in educational administration in 2005 at the University of Victoria. He has researched and published on the moral, political, emotional, and aesthetic dimensions of educational leadership, drawing theoretical inspiration from the Critical Theory tradition. Dr. Milley has significant professional experience in executive leadership development, policy research, and adult and workplace learning from a successful career in the Canadian federal public service and postsecondary institutions. He has researched, presented, and taught in numerous countries, including Brazil, China, the Netherlands, Singapore, and the United Kingdom.