

# “Deliverology” and Evaluation: A Tale of Two Worlds

---

Lisa Birch and Steve Jacob

Université Laval

**Abstract:** *In recent years, the new political governance, a partisan model that contributes to a permanent campaign, gained ground in public organizations. In this new context, “deliverology” is portrayed as an innovative method to help governments implement new policies and deliver on election promises. This article presents the similarities and differences that exist between “deliverology” and evaluation. Is deliverology really something new or is it another case of old wine in a new bottle? Is deliverology a substitute for or, instead, a complement to institutionalized evaluation? To what extent does new political governance (exemplified by deliverology and performance measurement) undermine evidence-based decision making? What is the value-added of deliverology? These questions are addressed through a critical reflection on deliverology and its value-added in Canada, where evaluation became institutionalized in many departments and agencies under the influence of results-based management, promoted by the advocates of new public management over four decades.*

**Keywords:** *deliverology, election promises, evaluation, new political governance, results-based management*

**Résumé :** *Depuis quelques années, la nouvelle gouvernance politique, un modèle de gestion partisane qui contribue à une campagne permanente, gagne en faveur dans les organisations publiques. Dans ce nouveau contexte, la « résultologie » est présentée comme une méthode innovatrice de mise en œuvre des politiques gouvernementales et de suivi de la réalisation des promesses électorales. Cet article présente les points communs et les différences qui existent entre la résultologie et l'évaluation. La résultologie est-elle vraiment quelque chose de nouveau ou est-ce seulement une autre manière de réinventer la roue? La résultologie est-elle un substitut ou plutôt un complément à l'évaluation qui s'est progressivement institutionnalisée dans l'appareil d'État? Dans quelle mesure la nouvelle gouvernance politique (incarquée par la résultologie et la gestion du rendement) vient-elle miner la prise de décisions éclairées par des données probantes? Quelle est la valeur ajoutée de la résultologie? Pour répondre à ces questions, le présent article propose une réflexion critique sur la résultologie et sa valeur ajoutée au Canada, où l'évaluation s'est institutionnalisée dans les ministères et les agences sous l'influence de la tendance de gestion axée sur*

**Corresponding author:** Lisa Birch, Département de Science politique, Université Laval, Pavillon Charles-De Koninck, local 3449, 1030, avenue des Sciences-Humaines, Québec, QC G1V 0A6; [lisa-maureen.birch@capp.ulaval.ca](mailto:lisa-maureen.birch@capp.ulaval.ca)

*les résultats, dont les partisans de la nouvelle gestion publique font la promotion depuis quelque 40 ans.*

**Mots clé :** *résultologie, promesses électorales, évaluation, nouvelle gouvernance politique, gestion axée sur les résultats*

Throughout his latest book, entitled “*What is Government Good at?: A Canadian Answer*,” Donald Savoie (2015), a well-known Canadian political scientist, is highly critical of the inverted 90 percent–10 percent culture<sup>1</sup> with a strong, structural fault line between ideas and implementation that has arisen with the advent of the permanent election campaign. He argues that 90% of policy resources should be dedicated to implementation and only 10% to ideas. The current inversion of these proportions has accentuated executive-level centralization and partisan dynamics in Canadian public administration in the context of an increasingly complex, globalized policy environment with rising pressures for higher transparency standards and ever more demanding media scrutiny. Chouinard and Milley (2015) raise concerns about how this twenty-first-century trend, coined the “new political governance” by Aucoin (2012), undermines the evaluation function and evidence-based approaches in policy formulation, generating two paradoxes: (1) the paradox of use, which shifts the focus toward symbolic utilization for technical tracking, as opposed to policy learning and social betterment; and (2) the paradox of accountability, which redirects attention toward democratic accountability and away from accountability for public management, leading to the politicization of public administration and an erosion of public-sector norms of neutrality, independence, and professionalism.

In *How to Run a Government so that Citizens Benefit and Taxpayers Don't Go Crazy*, Sir Michael Barber (2016), the founder of “deliverology,” criticizes what he calls “government by spasm,” which is caught up in present-focused crisis management, producing sound bites and announcements in anticipation of and reaction to “noise” from the conventional and social media agendas. He advocates instead for “government by routines,” which is focused on managing and driving its own delivery agenda, with regular reporting on progress and success stories to the media and the public.

Recourse to Barber’s “deliverology” as a tool for new political governance began with the Tony Blair government in the United Kingdom (1997–2007) and spread to Canada through the Dalton McGuinty government (Ontario, 2003–13) and now the Justin Trudeau government (Canada, 2015–present). However, there is a dearth of evidence on deliverology’s effectiveness as a “way to run a government” and little analysis of how it is differentiated from results- or performance-based management or from concomitant evaluation or monitoring (Richards, 2018; Richards, Gallo, & Kronick, 2017; Schacter, 2016). In the Canadian context, where the state progressively institutionalized evaluation and results-based approaches over four decades, one may wonder about the “value-added” of Barber’s deliverology and, as Schacter (2016) puts it, what problem deliverology is purporting to solve.

Although [Savoie \(2015\)](#) and [Barber \(2016\)](#) share a similar diagnosis of modern governance challenges and the need to focus on implementation (see note 1), they do not reference each other’s work or even the same literature on governance. Furthermore, where Savoie criticizes new political governance and private-sector-style performance management as harmful for implementation and service delivery in the public sector, deliverology embraces both the governing party’s reform agenda and private-sector-inspired performance management for optimal delivery, thereby becoming a tool for new political governance. These observations inspire the image of two worlds, each with its own referential frame and its own conclusions, which share a vision that government should realign to deliver whatever it is good at doing for the greater good of citizens. The small overlap between the two worlds remains a concern for improved decision making regarding policy implementation and for better policy outcomes. One domain is the “deliverologist’s” world and the other is the world of the evaluators, public administration scholars, and political scientists.

Through an overview of central ideas and criticisms of Barber’s deliverology, we situate deliverology relative to *ex ante*, concomitant, and *ex post* evaluation as well as results-based management. We then explore the controversial implementation of deliverology in Canada. Despite their common concerns for improved policy implementation, deliverology and evaluation appear largely as two worlds with different roots and evidence standards and with some overlapping zones, notably the use of performance indicators in evaluation and the aspiration for social betterment through effective policy delivery and meaningful public-sector reform. However, there are vast zones where these worlds diverge discursively, methodologically, and philosophically.

## WHAT IS BARBER’S DELIVEROLOGY?

Deliverology is a trademark created and owned by Delivery Associates, Michael Barber’s consultancy firm ([Delivery Associates, n.d.a](#)). Although Barber initially spoke of an emerging science of delivery, Delivery Associates market it as follows:

Deliverology® is our system for helping governments deliver meaningful results that will last. It is both a science and an art. The science is the routine of setting a target and then using data, technology, planning, monitoring and problem solving to achieve it. The art is the way you do it and how you behave—it needs focus, urgency, ambition, honesty and humility. ([Delivery Associates, n.d.b](#))

Deliverology is a global business that promotes a model of how to run a government drawn from business management and its founder’s practical experience in education and advising governments on delivering reforms. Barber relentlessly transformed his original practitioner experience acquired in the Blair years as head of the Prime Minister’s Delivery Unit into a tool for global career mobility through partnerships and networks involving high-profile management consultant firms such as McKinsey & Company with the McKinsey Global Institute, and

the Boston Consulting Group with its Centre for Public Impact (Brown, 2015; see also Centre for Public Impact, 2016, 2017, n.d.), and now his own Delivery Associates. This network includes international organizations like the OECD, the World Bank, the World Economic Forum (see Global Agenda Council on Education, 2016, for an example), and non-profit or philanthropic think tanks such as the Centre for American Progress and the Brookings Institute, as well as The Economist Intelligence Unit, the Institute for Public Policy Research, the Institute for Government, the Social Market Foundation, and NESTA, all in the United Kingdom. Deliverology is thus an example of a business in the lucrative, public-sector management reforms industry (see Manning, 2015, about this industry, and Manning & Watkins, 2013, for a comparison of deliverology to other approaches).

The main sources of inspiration for Barber's model of delivery, as shown by his bibliographies, span an eclectic mix, including (1) his own experiences as recounted in his publications, (2) historical accounts of country success, fixing failures, and trends in Western history, with works by Fukuyama, (3) biographies or autobiographies of leaders such as Winston Churchill, Theodore Roosevelt, Margaret Thatcher, and Tony Blair, (4) various publications by governments, consultancy groups, think tanks, and international organizations, (5) popular literature on great or successful companies, (6) accounts of delivery-unit successes in other countries such as Chile and Malaysia, and (7) books on public and private strategies (see reference lists in Barber, 2008, 2013, 2016, 2017; Barber, Donnelly, & Rizvi, 2013; Barber, Moffit, & Kihn, 2011a, 2011b). The preference for grey literature as opposed to peer-reviewed academic publications is summarized in the following statement:

synthesis is now provided by organisations outside universities—thinktanks, public agencies or consultancies, or by those who translate the synthesis into action.... [T]he most influential reports globally in the past five years have come from the Organisation for Economic Co-operation and Development (OECD), McKinsey or, recently, the Economist Intelligence Unit. (Barber et al., 2013, p. 17)

The academic literature in evaluation, public administration,<sup>2</sup> and political science is remarkably absent from Barber's reference lists. Conversely, references to "deliverology" *per se* are rare in evaluation and policy studies, where there is a vast literature grounded in scientific inquiry and practitioners' experiences after four decades of institutionalization and professionalization of evaluation in Canadian public administration (Canada, Treasury Board Secretariat, 2013–2014; Cummings & Gauthier, 2016; Dobell & Zussman, 2018; Jacob, 2006). Thus, the first divergence between evaluation and deliverology is rooted in their respective sources of knowledge and inspiration.

## DIVERGENT EVIDENCE STANDARDS

Flowing from this divergence is another one concerning evidence standards for policymakers. At first glance, deliverology seems at odds with discourses

on bringing scientific evidence back into policy. However, [Cairney and Oliver \(2017\)](#) remind us that scientists compete with other influential actors to access policymaking audiences. They argue that these audiences face two challenges. First, policymakers operate in complex, power-sharing, multi-level governance environments. Second, given this complexity and its inherent ambiguities, policy-makers act under conditions of bounded rationality and thus accept variations in evidence standards and hierarchies. Policymaking audiences can opt for evidence from implementation science, with its emphasis on gathering data on effectiveness, with attentiveness to evidence hierarchies that value systematic reviews and randomized controlled trials (RCTs), on scaling up best practices through replication, and on managing dosages of the “active ingredient” for change. They can also include other social science research as well as internal and external evaluations of policies and their implementation and outcomes ([Cantelli, Jacob, Genard, & de Visscher, 2006](#); [Nilsen, Ståhl, Roback, & Cairney, 2013](#)). They can choose evidence presented through storytelling approaches about indicators of effectiveness from practitioners and service users, narrations of experiences in case studies, and descriptions of key principles to be prioritized for service delivery (for an example, see *Results and Delivery—Lessons From Around the World*. [Canada School of Public Service, 2016a](#); see also [Delivery Associates, n.d.b](#)). In Canada, at the outset of his first mandate, Justin Trudeau, as Prime Minister of Canada, turned to Barber with his storytelling approach and more supple evidence hierarchy (see [Dobell & Zussman, 2018](#); [Puttick & Ludlow, 2013](#); [Shepherd, 2018](#)) instead of contracting to experts in public policy and public administration for guidance in implementing, rather than delivering, the Liberal Party of Canada’s platform. It is important to note that, for Barber, evidence that a policy is working lies in achieving a set of chosen targets set by the executive level of government, whereas for evaluators and scholars, the aspirational standards are set in light of the policy goals and require the production of valid and reliable data using scientific methods to determine the policy’s relevance, effectiveness, efficiency, impact, and sustainability in achieving positive intended outcomes.

## IS DELIVEROLOGY A MODEL OF POLICY CHANGE OR OF PLEDGE FULFILLMENT?

Let us now turn to the core elements of deliverology as a model of policy change and implementation and as a tool for reforming public-sector management. This model tells policymakers at the top to begin by answering five simple questions: (1) What are you trying to do? (2) How are you trying to do it? (3) How, at any given moment, will you know that you are on track? (4) If you are not on track, what are you going to do about it?, and (5) How can we help?, and to establish a routinized system requiring civil servants to report upwards on these same questions ([Barber, 2008](#); [Canada School of Public Service, 2016a](#); [Murekeyisoni, 2017](#)). The governing party answers these questions in reference to its election platform, giving a partisan orientation to the policy goals, programs, and reforms

to be implemented by civil servants. The implicit, normative assumption is that the governing party's policy choices and reforms are inherently correct and fully legitimate, thus potentially precluding *ex ante* assessments of their relevance, logic models, and evidence base.

For the next step in implementation, Barber advises politicians that they must achieve the four components of deliverology in order as follows: (1) develop a foundation for delivery, (2) understand the delivery challenge, (3) plan for delivery, and (4) drive delivery (Barber et al., 2011a, p. 237). The first step involves reviewing the current state of delivery and creating a specialized results and delivery unit at the apex of power that leads delivery. Located deep inside the executive level of government, the delivery unit is far above Savoie's (2015) fault line in the current 90% range dedicated to policy ideas and communications. Although it does not invert the problematic "90/10" split, it does bring a strong mandate to pursue implementation relentlessly down the delivery chain to produce, monitor, and report on meaningful outcomes using metrics that citizens can see and understand (Barber, 2016; Barber et al., 2011a, 2011b; Murekeyisoni, 2017). This specialized delivery unit is to lead the way to cultural change in government and throughout delivery chains such that the institutional buzzwords become "ambition, focus, clarity, urgency and irreversibility" (Barber et al., 2011a; see pp. 31–33).

The second step requires an assessment of past and present performance as well as the drivers of performance. Next, the government must proceed to establish a delivery plan comprising three key components: (1) the reform strategy, (2) the delivery chains, and (3) the targets and trajectories. The metrics must be clear, measurable, and meaningful to citizens and, especially, to politicians to help them "run a government" and usher in change efficiently. Finally, it is necessary to drive delivery by establishing routines for performance monitoring and problem-solving, and by maintaining the momentum. Thus, deliverology seems to address the bounded rationality problem of policymakers by proposing a simple plan: the relentless pursuit of a set of clear goals with the prioritization and tracking of key metrics and the delivery capacity to act quickly when needed.

Barber presented deliverology to the Canadian cabinet as a four-year pledge-fulfillment strategy: "Set the agenda and develop the approach" and "show early wins" (year 1), "establish the delivery discipline across government" and "drive for results" (year 2), "continue drive for results" and "plan delivery of the next parliament's top priorities" (year 3), and "manage through to success and run-up to the election" (year 4) (see Barber, 2016; Canada School of Public Service, 2016a). Journalists, commentators, and some academics tend to perceive deliverology as a pledge-fulfillment tool that helps the government to achieve measurable, concrete outcomes that respect election promises and keeps them accountable to citizens while improving governance through continuous monitoring of implementation (Arellano, 2016a, 2016b, 2016c; Castonguay, 2016; Lahey, 2017; Murekeyisoni, 2017; Wells, 2016a; Wherry, 2016, 2017; Zilio, 2016). Ironically, despite Barber's critique of short-term thinking and government by political communication,

deliverology promotes a management and communication strategy that fits with the electoral logic of the permanent campaign in representative democracy, where governing parties are always preparing for, recovering from, and engaging in elections (Couture & Jacob, 2019).

By defining the election promises as the deliverables and planning for their delivery over the electoral cycle, deliverology implicitly embraces the traditional model of electoral or democratic accountability known as promissory representation, whereby citizens grant a mandate to a political party to govern and hold the government accountable to keeping its election promises (Mansbridge, 2003). Research shows that, deliverology or not, the promise-keeping records at the federal level in Canada and provincially in Quebec are quite good (Pétry, 2014; Pétry & Birch, 2016, 2018; Pétry & Duval, 2015). The governing party seeks to satisfy voter preferences expressed by the election results. Recent work examining the performance of the Couillard government in Quebec (2014–2018) and the Trudeau government (2015–present) shows that while pledge fulfillment is an important aspect of governance between elections, much legislative and policy action is not linked directly to election promises but rather to agenda-setting dynamics for new issues and policy-cycle dynamics for existing policies and programs (Jacob, Birch, Pétry, & Baby-Bouchard, 2018; Birch, Jacob & Baby-Bouchard, 2019). Furthermore, in a pioneering study, Werner (2018) discovered that when given a choice between democratic representation models, fewer than one in ten respondents preferred promissory representation through faithful promise keeping, while half chose the trustee or anticipatory model of governing for the common good (see also Mansbridge, 2003, for a discussion of all representation models). Without a mutually exclusive choice between models (i.e., with separate questions about each model), respondents gave each representation model high importance rankings. Werner concludes that voters “want it all”—governing parties that are promise keepers and trustees of the common good and that show responsiveness to public opinion. In light of these findings regarding democratic accountability, it is necessary to ask what the value-added of deliverology is, as a way to implement election promises deemed to be of inherent worth, when governments are already quite good at promise keeping and when citizens may expect more than promise keeping.

## HOW DOES DELIVEROLOGY COMPARE TO EVALUATION AND RESULTS-BASED MANAGEMENT?

On the surface, deliverology seems to include a form of monitoring or concomitant evaluation (as opposed to *ex ante* or *ex post* evaluation), with its emphasis on tracking and targeting key metrics. These metrics are similar to results-based targets in public-sector management. Tweedie (2018), a policy developer and strategist in the government of Canada, argues that deliverology is really a rebranding of results-based management. Deliverology supports recourse to dashboards, but with a small set of specific targets deemed to be meaningful for citizens.

Both evaluation and deliverology share an ambition to ensure social betterment through the development, design, and implementation of effective policies and programs that yield meaningful outcomes. However, deliverology focuses on implementing the governing party's vision of betterment via a delivery unit, while the multiple evaluation units dispersed throughout government agencies question the coherence between policy aims, implementation, and outcomes, their appropriateness, and the soundness of underlying logic models. As [Table 1](#) shows, there are many substantive distinctions between the two worlds. Additional differences concern the vocabulary and scope of modern theory-based evaluations that attend to both the intended and unintended outcomes, using a variety of qualitative and quantitative data sources rather than a shortlist of key metrics.

### **WHAT ARE THE CRITICISMS OF DELIVEROLOGY AND ITS CONSEQUENCES FOR EVALUATION AND EVIDENCE-INFORMED POLICYMAKING?**

Positive perspectives on deliverology start from Barber's claims that it enhances public-sector management, performance, and accountability to better serve citizens, regardless of the conditions at the outset of adopting deliverology. [Barber \(2016\)](#) and [Castonguay \(2016\)](#) suggest that deliverology can alleviate taxpayer dissatisfaction and improve confidence in government. However, most of these claims are prospective opinions that have not been verified empirically through a rigorous independent evaluation process.<sup>3</sup> Robert [Lahey \(2017, p. 7\)](#), a Canadian evaluation expert and practitioner turned consultant, who was the founding head of Canada's Centre of Excellence for Evaluation, noted, "On the positive side, deliverology, given its link to political power, generally brings with it authority, resources, flexibility and a striving for provision of timely advice and quick turnaround (i.e. a sense of urgency that can potentially cut through bureaucratic roadblocks to action)."

In one of the rare independent case studies of the concept, [Murekeyisoni \(2017\)](#) finds positive effects in her assessment of deliverology applied to Ontario's education reform in the McGuinty years. She found that deliverology facilitated efficient management and coordination while ensuring a citizen-centered, results-based approach to educational reform. She further noted that it helped address resistance to change in the delivery chain. However, this success reflects adaptations of the deliverology model to engage actors collaboratively in defining goals and targets, rather than a standard top-down delivery pattern, and to grant flexibility to actors to adapt to change.

Many negative criticisms situate deliverology as a variant of results-based management and refer to research findings about the pitfalls of the application of private-sector-inspired reforms with performance targets in the public sector ([Hood & Dixon, 2015a, 2015b](#); [Manning, 2015](#)). For example, deliverology is criticized for accentuating the hierarchical concentration of power within the Prime Minister's Office (PMO), to the detriment of accountability



**Table 1.** Contrasting deliverology and evaluation

Comparison point	Deliverology	Evaluation
Purpose	Promotion and implementation of delivery units as a method of ensuring policy and reform delivery to improve government functioning to benefit the lives of citizens	Production of valid and reliable knowledge to decision makers about programs and organizations with a view to social and organizational betterment
Target audience for reporting	Prime Minister, Cabinet, and the electorate	Managers and executive decision makers within government
Type of accountability	External accountability	Internal accountability
Focal point(s)	Fulfillment of government's agenda, election promises, and public-sector reforms	Fulfillment of stakeholders' needs for data about relevance, accessibility, comprehensiveness, integration, fulfillment of objectives, effectiveness, impact, cost, efficiency, and sustainability (OECD, 2010; Patton, 1997).
Evidence base—main sources	Practitioner experiences and stories of successful case studies about implementation	Scientific and evaluative knowledge gathered using social research methods to determine how and why what works and does not work for whom in different contexts
Perspective on evidence and policy	Discursive support of the idea of "evidence-inspired" policy	Implicit and explicit commitment to the idea of evidence-informed policy
Approach to data analysis	Tracking of selected metrics for reaching targets for each government priority to assure "relentless implementation"	Analyzing multiple quantitative and qualitative data sources to ensure a complete picture
Approach to identifying indicators	Considering metrics chosen by the executive with the Delivery unit as inherently appropriate	Using logic theories to identify relevant standards or performance indicators and to focus on overall intended outcomes
Time horizon	Immediate and short-term assessments during an election mandate	Multiple assessments of outcomes and impacts over time, with a focus on the medium- and long-term impact
Adherence to standards	Ensuring transparency about delivery of the governing party's pledges	Respecting formal program evaluation standards (Joint Committee on Standards for Educational Evaluation) and ethics principles

Note: Information adapted from Barber et al. (2011a, 2011b), Canadian Evaluation Society (2014)

mechanisms such as institutionalized evaluation in the public service and reporting to Parliament. Top-down command-and-control approaches tend to demotivate public-service workers by failing to recognize their existing expertise and by considering them as obstacles to policy reform rather than partners in delivery. Empirical work analyzing 20 agencies operating with performance contracts in Quebec revealed that better outcomes arose when decision-authority and financial controls were decentralized, not centralized (Aubert & Bourdeau, 2012), which is in keeping with Murekeysoni's (2017) comments on the adaptations of deliverology in Ontario. In contrast to Murekeysoni's and Barber's own claims about the effectiveness of deliverology for Ontario's education reforms, Sattler (2012) found that these education-governance reforms were rooted in a neoliberal accountability and performance reform agenda starting in 1990 that was implemented through a long-term, incremental, and "messy" process, with no evaluation of the ultimate impact on outcomes for students. Furthermore, some UK case studies also suggest a continuum with pre-existing new public management reforms and expose the limits to deliverology due to short-sighted targets, missed delivery problems, the complexity of policy implementation, and the challenges of changing organizations and target population behaviors in durable ways (Smith, Richards, Geddes, & Mathers, 2011; Smithers, 2001).

A second line of criticism concerns the targets themselves, pressures to reach targets, and unintended, perverse, or short-lived effects. When targets lack legitimacy among practitioners, focus on a narrow set of metrics to the exclusion of others, and include mixed incentives attached to employee performance assessment or to program funding, they can undermine results and outcomes, yielding perverse, unintended results by demotivating and devaluing practitioners, creating a climate of fear about the consequences of missing targets, and by placing them in ethical dilemmas (Arnaboldi, Lapsley, & Steccolini, 2015; Ayers, 2013; Jacob, 2009a, 2009b; Jacob, Miller-Pelletier, & Birch, 2017; Murekeysoni, 2017; Poister, 2003). By focusing on chosen metrics, deliverology may sacrifice the systematic analysis of unintended, even perverse consequences that will arise if actors down the delivery chain find incentives to "present selective data," "game the metrics," and "distort or manipulate the findings" (Bevan, 2012; Bevan & Hood, 2006; Fisher & Downes, 2008; Hood, 2006, 2010, 2012; Schacter, 2002a, 2002b, 2006, 2011, 2016; Welsh & Ordóñez, 2014). The harshest and most sarcastic criticism of deliverology comes from John Seddon (2008, 2013), who argues using examples from deliverology under the Blair government that it is an oversimplified form of "Mickey Mouse command and control" that believes in "reform by targets" but misses the desired policy outcomes. He concludes that deliverology results in a dysfunctional, distorted system where demoralized, cynical service providers "manage all the wrong things," by focusing on costs rather than values and on gaming targets rather than achieving higher outcomes, which leads to increased costs to support gaming strategies, with perverse effects on outcomes especially in the health-care and education sectors (see CFA Local 1983 [2010a, 2010b] for videos of Seddon's speeches condemning deliverology). To the extent

that deliverology is “old wine in a new bottle,” these problems with results-based approaches are relevant.

Another consequence of results-based approaches is the trend toward using evaluation expertise in Canadian public administration to develop, track, and report on a limited set of metrics, which diminishes the time and resources allocated to conducting all-encompassing evaluations that ensure accountability and learning for outcomes and do not just target achievements and reporting up the internal hierarchy (Savoie, 2015). Politically driven and externally oriented accountability mechanisms embedded in performance regimes seem to miss the mark for measuring performance and for creating positive incentive systems for enhanced service delivery and policy learning, especially when policies entail complex tasks that are difficult to capture in a limited range of indicators (Jakobsen, Baekgaard, Moynihan, & van Loon, 2018). Already, without the addition of deliverology, this narrow approach undermines the role and production of *ex post* evaluations as a tool for policy learning and improving outcomes (Shepherd, 2018). When delivery units are created within departments in parallel with existing audit and evaluations departments, multiple tensions arise between the functions of delivery, auditing, and evaluation. Delivery units, because of their direct links to the PCO and their preoccupation with results indicators, further institutionalize a form of new political governance that may squeeze audit and evaluation functions. For civil servants, ethical tensions can emerge between the public-sector ethos of neutrality and the perceived partisan prerogatives of delivery units, especially when the line between public service in the government and public service under the governing party becomes fuzzy.

Furthermore, deliverology targets are more limited in range and depth than the indicators of program relevance, effectiveness, and outcomes that evaluators would choose or that may be meaningful to citizens. When deliverology targets are met while other more significant indicators are not, this may actually undermine legitimacy, as citizens perceive a gap between the successful attainment of a target and their own experiences as the beneficiaries of a public service. When the broader range of evaluation criteria and indicators is ignored, it is likely to take more time and be more tedious to improve delivery, beneficiary satisfaction, and desired long-term outcomes. Although command-and-control approaches to delivery may yield short-term results, long-term change requires modifications in organizational culture and behaviour. This may stall as the next election nears because the delivery concerns of politicians metamorphose into a focus on the campaign ahead, and those of civil servants anticipate those of the next government.

Although Barber recognizes the growing criticism of the effectiveness of delivery units, given their variable results, his response is as follows:

This critique is important, but is not completely fair, as the idea of a Delivery Unit has become a victim of its own success. When something is fashionable, lots of people rush to do it, and so impact inevitably varies. As a result, many Delivery Units have the label, but not much else. They belong to a large and growing group of “DINOs”—Delivery Units in Name Only. (Delivery Associates, n.d.b)

A third line of criticism tackles the question of the value-added of deliverology from the perspective of the practitioners of public policy, management, and evaluation as well as of scholars and students of public policy and management. Essentially, deliverology just repackages old ideas, especially but not exclusively about performance management. Many of these ideas, which are standard fare in introductory “101” courses on public policy, administration, or project management, are portrayed as innovative ones by advocates of deliverology. This repackaging is well illustrated in two reports by the Centre for Public Impact, which is co-chaired by Michael Barber, namely “The Public Impact Fundamentals Report: Helping Governments Progress from Idea to Impact” (2016) and “A Rubric for Assessing Public Impact: Applying the Public Fundamentals to the Real World” (2017). These reports present a self-proclaimed “new maximization framework” to close the gap between what governments achieve for their citizens and what they could deliver, a “checklist for practitioners” based on the three public fundamentals (legitimacy, policy, and action), and a rubric for determining what sources of information (such as organization publications, initiative blueprint, initiative implementation plan, speeches from politicians, party manifestos, and public opinion research, plus articles, case studies, and reports) are needed to evaluate the fundamentals. The promotion of these simple, repackaged frameworks and checklists as “new” to professionals who have been applying them and more for some time in their departments errs by ignoring their pre-existing knowledge and expertise. This can foster cynicism and demotivation rather than enthusiasm for the government’s reform agenda. This is more likely when civil servants perceive delivery units as restructuring the space for policy conversations and the roles of existing internal accountability units to their disadvantage or perceive new conflicts between goals, short-sighted targets, and outcomes. Furthermore, according to [Jakobsen et al. \(2018\)](#), external accountability regimes (of which deliverology is one) that are not grounded scientifically and empirically tend to “crowd out,” rather than stimulate, the intrinsic motivation of service providers because of mismatched external incentives and performance indicators, weak consideration of the complexity of policy tasks, and implicit disregard of the importance of intrinsically motivated, autonomous professionals in service management and delivery.

Finally, one may ask whether the “old wine in new bottles” remains, nonetheless, of timeless relevancy or is past its expiry date. This is a question that those who have been promoting the deliverology model relentlessly since the Blair years do not seem to entertain. By contrast, this is the quintessential question for evaluators, public administration scholars, and political scientists. Their theoretical reflections, empirical research, and practitioner experiences regarding results-based management, its organizational and whole-of-government impact, and its relationship to evaluation have supported a trend favoring “*new* wine in new bottles.” This trend goes beyond earlier, simplified, top-down approaches to results-based management to recognize the importance of dialogues between those who make the policy decisions and those who assure implementation. These dialogues involve conversations between agents and principals about performance, the

interface between performance management and evaluation, designing performance systems to encourage organizational learning and effectiveness, policy task complexity, the appropriate indicators, employee motivation, incentive mechanisms as an antidote to cynicism, and the optimal conditions for policy as well as organizational learning (e.g., [Jakobsen et al., 2018](#); [Laihonen & Mäntylä, 2017](#); [Moynihan, 2005, 2008a, 2008b, 2010](#)).

## WHAT ABOUT THE IMPLEMENTATION OF DELIVEROLOGY IN CANADA?

Deliverology, like any public management reform, is never implemented in a vacuum. For this reason, let us begin by establishing the context in which it came to Canada through a brief review of historical developments in Canadian public policy and administration. From the 1970s onward, Canada began experimenting with results-based approaches. The country undertook yet another initiative<sup>4</sup> on “re-thinking government” in the mid-1990s and implemented routines for evaluation, results-based management, regulatory impact assessment, citizen engagement and consultation, stakeholder participation, the production and use of public opinion research, evaluation, and scientific evidence in policy. Although some of these routines were under pressure during the Harper years, they did not disappear from the public-sector landscape. Rather, the Canadian state honed and fine-tuned the infrastructure for monitoring and evaluation and the requirements for the utilization of evaluation results that were instituted with the Treasury Board’s Results for Canadians in 2000 and reinforced with mandatory performance frameworks at the program and organizational levels ([Lahey, 2017](#); [Lahey & Nielsen, 2013](#); [Shepherd, 2018](#)). Since 2013, long before deliverology and the mandate letter tracker, the Treasury Board already published online the GC InfoBase, which tracks a variety of indicators and government results ([Canada, Treasury Board Secretariat, n.d.](#)). The 2009 *Evaluation Policy* ([Canada, Treasury Board Secretariat, 2009](#)) that was replaced by the *Policy on Results* ([Canada, Treasury Board Secretariat, 2010](#)) already talked about producing timely, neutral evaluation evidence and managing for results.

In many ways, Canadian public servants have been working within results-based management and accountability frameworks since the 1970s, long before the Trudeau government contracted Barber’s services ([Curran, 2016](#); [Lahey, 2017](#); [Lahey & Nielsen, 2013](#); [Tweedie, 2018](#)). In fact, [Tweedie \(2018\)](#) asserts,

We are open, accountable, and transparent. We use plain language and consult with our stakeholders. We have strategic plans, operational plans and implementation plans. We use SMART indicators and ensure that relevance and performance (value for money) drive our organisations to deliver results of benefit to Canadians. We do risk assessments, audits, and evaluations according to clear, published criteria. Modern public policy governance? Nailed it.

However, both [Tweedie \(2018\)](#) and [Savoie \(2015\)](#) point out that the Auditor General of Canada remains critical of public services where the implementation of the

results-based approach resulted in agencies focusing on internal accountability chains, thus neglecting the citizens' perspective and the unintended consequences of the policy in their performance reports (Auditor General of Canada, 2002, 2004, 2005, 2017). Lahey and Nielsen (2013) also note that the Auditor General supports monitoring and evaluation practices and uptake and frequently calls for more substantive evaluation of overall program effectiveness and quality of results. They also recognize the persistent difficulties the Canadian system experiences in "establishing the right balance" (p. 49) between using performance measurement from monitoring and evaluation for accountability versus policy learning. Given this historical backdrop, deliverology, with its focus on delivering the governing party's agenda and on defining success in terms of reaching targets, may well exacerbate these difficulties by reinforcing this pre-existing trend to favour performance measurement use for control rather than for improving operations through policy learning. Let us now turn to the implementation of deliverology in Canada at the federal level.

On December 23, 2015, Prime Minister Trudeau named Matthew Mendelsohn, former advisor and deputy minister in Ontario's McGuinty government, as Deputy Secretary to the Cabinet (Results and Delivery) in the Privy Council Office. At the first Liberal cabinet retreat in St. Andrews, New Brunswick, in January 2016, deliverology officially appeared on the federal political landscape (Dyer, 2016a). Sir Michael Barber from Delivery Associates, under contract for two years at \$217,000 (Castonguay, 2016), Dominic Barton, CEO from McKinsey Consulting, working pro bono (Geddes, 2016), and their ideas became special guests at this meeting and at subsequent federal cabinet retreats in Kananaskis, Alberta (April 2016), Sudbury, Ontario (August 2016), St. John's, Newfoundland and Labrador (September 2017), and London, Ontario (January 2018). Through the Results and Delivery Unit in the Privy Council Office, Barber's ideas became embedded in the Trudeau government, which created a cabinet committee on "Agenda and Results," chaired by the Prime Minister, which would become the "Agenda, Results and Communication" committee by August 2016 (Wells, 2016b) and subsume the cabinet committee on open government and parliamentary affairs in August 2018. The Canada School of Public Service offered videos featuring Barber (Canada School of Public Service, 2016a, 2016b), and the Institute on Governance (2018) offered training packages on deliverology. The Trudeau government deployed considerable effort at the apex of power to show a commitment to effect a deliverology culture shift from the top of the Canadian government down the delivery chains.

By July 1, 2016, the government had introduced the *Policy on Results* and the *Directive on Results* to replace the *Policy on Reporting of Federal Institutions and Corporate Interests to Treasury Board of Canada Secretariat* (2007), the *Policy on Evaluation* (2009), and the *Policy on Management, Resources and Results Structures* (2010) (Canada, Treasury Board Secretariat, 2016a, 2016b). Under the new policy, the evaluation function remains embedded in the Canadian state, with some new requirements. The *Policy on Results* explicitly links evaluation with performance measurement semantically and structurally through the Performance

Measurement and Evaluation Committee. Departments began creating new positions for a Chief Results and Delivery Officer and a Chief Data Officer.

Then, in November 2017, the Trudeau government launched the ultimate deliverology tool: the Mandate Letter Tracker ([Canada, Privy Council Office, 2018a, 2018b](#); [MacCharles, 2017](#)). In January 2018, Matthew Mendelsohn, head of Canada’s delivery unit, presented deliverology as a “data-driven, results-focused approach” to public-sector management that included a space for evidence and evaluation, despite the challenges of “obtaining timely data for metrics on pledge fulfillment” (Mendelsohn, as interviewed by [Haws, 2018](#)). He seems to suggest that the value-added of the Mandate Tracker is to provide greater transparency about pledge tracking, a task that previous governments did behind closed doors, and to focus executive-level attention on implementation through “ongoing evaluation” as opposed to conventional *ex post* evaluations. In contrast, political scientists express serious concerns that deliverology encourages deeper centralization at the executive level with special communication channels and additional resources for delivery units (Alex Marland and Anna Esselment, as interviewed by [Ryckewaert, 2016](#)).

From the start, the Canadian media responded to these developments in federal governance with skepticism, calling Barber a “guru” ([Arellano, 2016c](#); [Dyer, 2016b](#); [McGregor, 2016](#)), the Prime Minister a “devotee” ([Wherry, 2016](#)), and deliverology as the latest “fad and fashion” ([French, 2016](#)) or a “punchline” ([Radwanski, 2017](#)), with growing criticism ([Wherry, 2017](#)) that switched to sarcasm when denied an interview with Barber on broken election promises and the use of “deliverology for relentless implementation” ([Bimman, 2018](#)). Responding to an article in *L’Actualité* ([Castonguay, 2016](#)), the Canadian Evaluation Society reminded readers of the 40 years of evaluation practice in Canada and the institutionalization of evaluation in the federal administration starting in 1977 ([Cummins & Gauthier, 2016](#)). At a press conference to release fall reports, Michael Ferguson, the [Auditor General of Canada \(2017\)](#), signalled that there is a delivery gap between the whole-of-government level and its messages about citizen-centric service delivery. When queried about the mandate tracker, Ferguson remarked that the limited scope of the mandate tracker with its focus on election promises missed the fact that “there’s more to government than big-talk mandate letters for ministers” (as cited by [Delacourt, 2017](#)). In many ways, the “Mandate Letter Tracker” website seems to be a new tool for political communications and packaging the government’s progress (or lack of progress) toward achieving its election commitments as reformulated and adjusted in the mandate letters.<sup>5</sup>

The arrival of deliverology in Canada sparked a friendly debate between [Mark Schacter \(2016\)](#), an Ottawa-based consultant and author of a critical essay entitled “Does ‘Deliverology’ Matter?,” and [Steve Montague \(2016\)](#), an adjunct professor and consultant who offered “A Response to Does ‘Deliverology’ Matter?: Targets Can Work If Done Properly.” Although Montague recognized many of Schacter’s criticisms, he expressed reservations about skepticism toward targets. He argued for the “responsible” development and management of targets

over the longer policy cycle in an iterative process that engages stakeholders and decision makers and contributes to policy learning. Thus, Montague implicitly embraced the “new wine” approach based on performance dialogues for learning and explicitly hoped that the Canadian brand of deliverology would engage in “responsible” targeting.

## CONCLUSION

This tale of two worlds traces the portrait of deliverology as a business and a model of implementation that repackages and rebrands old ideas linked to results-based management and democratic accountability as pledge fulfillment, while assuming that current governance systems are not delivering and therefore need a deliverology fix. It questions the innovativeness of deliverology generally and highlights some criticisms of this model, especially regarding the possibility that deliverology may usurp evaluative capacity. It finds limited value-added, since recourse to deliverology in Canada ignores both the strong record of pledge fulfillment and the particularities of Canada’s complex multi-level governance context. Although deliverology does bring a focus on implementation to the apex of power, its relevance and effectiveness seem somewhat elusive in a context where the state has already institutionalized evaluation and results-based approaches and where new political governance undermines frontline delivery capacity via the inverted “90/10” split. The effectiveness of deliverology as a “way to run a government” has never been subjected to an independent, whole-of-government evaluation, and, if it were, it would be difficult to disentangle outcomes owing to deliverology from those owing to other factors in a context of institutionalized evaluation and monitoring systems. Originally a “tongue-in-cheek” term coined by Barber’s critics (*The Economist*, 2015), deliverology became Barber’s globalized consulting enterprise. Barber’s deliverology has an aura of scientificity because of recourse to target metrics, and an appealing message because of its practical, business-like approach to complex governance challenges associated with public sector reforms.

Deliverology is light on science and heavy on narrative, with flexible evidence standards that favour “best practice stories” and narrower “targeted metrics,” whereas evaluation and social research are science heavy and narrative light, with stricter evidence standards, assessment criteria, and multiple metric points tailored to specific purposes. Each world has its own value orientation. Barber relies on business-inspired thinking about running a government that inherently values private-sector performance management. His model is shaped by the ideological orientations of the self-proclaimed, “small ‘l’ liberal” and “centre-right” think tanks and organizations that Barber refers to in his works. By contrast, evaluators and policy researchers are grounded in scientific thinking, leading them to value systematic recourse to theories, research methods, and empirical findings to address policy questions, to assess implementation outcomes, and to determine the effects of programs and policies, as compared to the “counterfactual”—that is, what would have occurred had the policy or program not been implemented,



or if they had been implemented differently. To them, many of Barber’s sources are “grey literature.” To Barber, the work of evaluators and policy scientists is too abstract and lacks synthesis. He cares about helping governments to deliver their election promises and reforms, whatever they may be, abstracted from any critical analysis of the reform agenda itself, the need for structural change within the public administration system, or the current pledge-delivery capacity of the system. By contrast, evaluators and political scientists care about the relevancy, appropriateness, and impact of such reforms—that is, their ability to achieve socially desirable outcomes with target populations in meaningful, cost-effective ways, and their ability to enhance public management and service delivery. Where evaluators aspire to facilitate conversations about “good” commitments and outcome congruence with policy goals, “deliverologists” assume party commitments are “good” and seek conversations about implementation and progress toward achieving set targets.

Perhaps Barber can learn from evaluators, public administration scholars, and political scientists about the importance of context, evidence of the perverse effects of governance by targets, evidence of the failures of private-sector-inspired reforms on public-sector capacity and employee motivation, and, most important, the role of more comprehensive evaluations of reforms, programs, and policies in deep policy learning and delivery of better outcomes. If nothing else, perhaps evaluators, public administration scholars, and political scientists should take note of the importance of “storytelling” and “packaging” as a mode of knowledge diffusion for uptake by policymakers. Barber’s success in global networking to diffuse his “Deliverology” model shows that storytelling about the effectiveness of this governance model, even without solid, science-based evidence, is more appealing to policymakers who say they support “evidence-informed” policy and management but who struggle with bounded rationality in complex, challenging policy environments.

## ACKNOWLEDGEMENTS

We wish to thank Raphaëlle Jacques for her contributions to the reference list. We acknowledge funding from The Fonds de recherche du Québec—Société et culture (FRQSC) for our research team.

## NOTES

- 1 [Savoie \(2015\)](#) asserts that “To those inside, particularly where it matters increasingly in central agencies and in the Prime Minister’s Office (PMO), government is 90 per cent ideas and 10 per cent implementation. To those outside, government should be 90 per cent implementation and 10 per cent ideas.... There is a fault line in government that separates those responsible for generating new policies and managing the blame game from those responsible for implementation. That explains why government was not as good as it once was in delivering programs and services” (pp. 14–15). Like Savoie, [Michael Barber \(2016, p. 36\)](#) advises politicians that “policy is 10 per cent and implementation is 90 per

cent.” Oddly, neither Savoie nor Barber references the other’s work, nor do they reference similar literature, despite their shared diagnosis of the 90/10 split and their criticism of the short-sightedness of “government by political communications.”

- 2 The exception is a reference to Osbourne and Gaebler’s famous book *Reinventing Government: How the Entrepreneurial Spirit Is Transforming the Public Sector* (1993), which emphasizes market-oriented, results-oriented governments that “steer” rather than “row.”
- 3 We began our research efforts and analysis of “deliverology” in 2016 at the same time as others (see Richards, 2018; Richards et al., 2017; Schacter, 2016). Our findings corroborate one another’s. We add that much of the positive literature on “deliverology” in cyberspace is linked to Barber or to one of the consulting firms with which he is or was associated.
- 4 Since the Glassco commission of 1962–63, the Canadian government has experienced multiple iterative processes that have led to the institutionalization of evaluation and shifts toward evidence-based decision making.
- 5 Governing parties in Canada have used similar tools to track their promise delivery, without calling it “deliverology.” See the Campbell government’s “Five Great Goals for a Golden Decade” in B.C., the Stelmach government’s “Measuring Up: Progress Report on the Business Plan of the Alberta Government,” the Charest government’s short-lived tracking tool or the traffic light system of the Marois government in Quebec, and the federal government’s “Departmental Results Framework and Inventory Program.”

## REFERENCES

- Arellano, N. (2016a). Can deliverology help solve education, health problems in remote areas? *Canadian Government Executive*. Retrieved from <http://canadiangovernmentexecutive.ca/can-deliverology-help-solve-education-health-problems/>
- Arellano, N. (2016b). Trudeau’s five deliverology targets. *Canadian Government Executive*. Retrieved from <http://canadiangovernmentexecutive.ca/trudeaus-five-deliverology-targets/>
- Arellano, N. (2016c). Trudeau’s “deliverology” guy. *Canadian Government Executive*. Retrieved from <http://canadiangovernmentexecutive.ca/trudeaus-deliverology-guy/>
- Arnaboldi, M., Lapsley, I., & Steccolini, I. (2015). Performance management in the public sector: The ultimate challenge. *Financial Accountability & Management*, 31(1), 1–22. <https://doi.org/10.1111/faam.12049>
- Aubert, B. A., & Bourdeau, S. (2012). Public sector performance and decentralization of decision rights. *Canadian Public Administration*, 55(4), 575–598. <https://doi.org/10.1111/j.1754-7121.2012.00238.x>
- Aucoin, P. (2012). “New Political Governance” in Westminster systems: Impartial public administration and management performance at risk. *Governance: An International Journal of Policy, Administration and Institutions*, 25(2), 177–199. <https://doi.org/10.1111/j.1468-0491.2012.01569.x>
- Auditor General of Canada. (2002). Chapter 6: A model for rating department performance reports. In *2002 April Report of the Auditor General of Canada*. Retrieved from [https://web.archive.org/web/20190618194158/http://www.oag-bvg.gc.ca/internet/English/parl\\_oag\\_200204\\_e\\_1133.html](https://web.archive.org/web/20190618194158/http://www.oag-bvg.gc.ca/internet/English/parl_oag_200204_e_1133.html)

- Auditor General of Canada. (2004). Chapter 7—Managing government: A study of the role of the Treasury Board and its Secretariat. In *2004 March Report of the Auditor General of Canada*. Retrieved from [http://www.oag-bvg.gc.ca/internet/English/parl\\_oag\\_200403\\_07\\_e\\_14899.html#ch7hd4b](http://www.oag-bvg.gc.ca/internet/English/parl_oag_200403_07_e_14899.html#ch7hd4b)
- Auditor General of Canada. (2005). Chapter 5: Rating selected department performance reports. In *2005 Report of the Auditor General of Canada*. Retrieved from [http://www.oag-bvg.gc.ca/internet/English/parl\\_oag\\_200504\\_05\\_e\\_14936.html](http://www.oag-bvg.gc.ca/internet/English/parl_oag_200504_05_e_14936.html)
- Auditor General of Canada. (2017, November 21). *Auditor General's opening statement—2017 Fall reports press conference*. Retrieved from [http://www.oag-bvg.gc.ca/internet/English/osm\\_20171121\\_e\\_42705.html](http://www.oag-bvg.gc.ca/internet/English/osm_20171121_e_42705.html)
- Ayers, R. S. (2013). Building goal alignment in federal agencies' performance appraisal programs. *Public Personnel Management*, 42(4), 495–520. <https://doi.org/10.1177/0091026013496077>
- Barber, M. (2008). *An instruction to deliver: Tony Blair, the public services and the challenge of delivery*. London, England: Politicos.
- Barber, M. (2013). *The good news from Pakistan*. Retrieved from <https://web.archive.org/web/20180823130848/http://www.reform.uk/publication/the-good-news-from-pakistan/>
- Barber, M. (2016). *How to run a government so that citizens benefit and taxpayers don't go crazy*. London, England: Penguin Books.
- Barber, M. (2017). *How to deliver improved outcomes for schools*. Retrieved from <https://www.wise-qatar.org/2017-wise-research-improved-outcomes-school-systems/>
- Barber, M., Donnelly, K., & Rizvi, S. (2013). *An avalanche is coming: Higher education and the revolution ahead*. London, England: Institute for Public Policy Research. Retrieved from <http://www.avalancheiscoming.com/>
- Barber, M., Moffit, A., & Kihn, P. (2011a). *Deliverology 101: A field guide for educational leaders*. Thousand Oaks, CA: Sage.
- Barber, M., Moffit, A., & Kihn, P. (2011b). *Deliverology: From idea to implementation*. Washington, DC: McKinsey and Co.
- Bevan, G. (2012). *The challenge of designing “good enough” performance measures & results framework*. London, England: London School of Economics.
- Bevan, G., & Hood, C. (2006). What's measured is what matters: Targets and gaming in the English public health care system. *Public Administration*, 84(3), 517–538. <https://doi.org/10.1111/j.1467-9299.2006.00600.x>
- Birch, L., Jacob, S., & Baby-Bouchard, A. (2019). The Trudeau government's legislative agenda: Election promises and a dual mandate. In L. Birch & F. Pétry (Eds.), *Assessing Justin Trudeau's Liberal government: 353 promises and a mandate for change* (pp. 27–42). Quebec, QC: Presses de l'Université Laval.
- Bimman, A. (2018, January 12). Analysis: Trudeau cabinet's special guests. *Global News*. Retrieved from <https://globalnews.ca/news/3962858/trudeau-cabinet-retreat-guests/>
- Brown, A. (2015, March 10). *Helping governments govern*. Retrieved from <https://www.bcg.com/publications/2015/public-sector-helping-governments-govern.aspx>

- Cairney, P., & Oliver, K. (2017). Evidence-based policymaking is not like evidence-based medicine, so how far should you go to bridge the divide between evidence and policy? *Health System Policy and Research*, 15(1), 35. <https://doi.org/10.1186/s12961-017-0192-x>
- Canada, Privy Council Office. (2018a). *Mandate letter tracker: Delivering results for Canadians*. Retrieved from <https://www.canada.ca/en/privy-council/campaigns/mandate-tracker-results-canadians.html>
- Canada, Privy Council Office. (2018b). *The mandate letter tracker and the results and delivery approach*. Retrieved from <https://www.canada.ca/en/privy-council/services/results-delivery-unit.html>
- Canada School of Public Service. (2016a). *Results and delivery—Lessons from around the world* [Video file]. Retrieved from <https://www.youtube.com/watch?v=jUJIDAy6KbY>
- Canada School of Public Service. (2016b). *Sir Michael Barber: Results and delivery unit: Lessons for Canada*. Matthew Mendelsohn interviews Michael Barber [Video file]. Retrieved from <https://www.youtube.com/watch?v=JogXHgPXhVs>
- Canada, Treasury Board Secretariat. (2007). *Policy on reporting of federal institutions and corporate interests to Treasury Board Secretariat*. Retrieved from <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=12622>
- Canada, Treasury Board Secretariat. (2009). *Policy on evaluation*. Retrieved from <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=15024>
- Canada, Treasury Board Secretariat. (2010). *Policy on management, resources and results structures*. Retrieved from <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=18218>
- Canada, Treasury Board Secretariat. (2013–2014). *Evaluation of the 2009 policy on evaluation*. Retrieved from <https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/centre-excellence-evaluation/evaluation-2009-policy-evaluation.html>
- Canada, Treasury Board Secretariat. (2016a). *Policy on results*. Retrieved from <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=31300>
- Canada, Treasury Board Secretariat. (2016b). *Directive on results*. Retrieved from <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=31306>
- Canada, Treasury Board Secretariat. (n.d.). *Infographic for government*. Retrieved from <https://www.tbs-sct.gc.ca/ems-sgd/edb-bdd/index-eng.html#orgs/gov/gov/infograph/results>
- Canadian Evaluation Society. (2014). *What is evaluation?* Retrieved from <https://evaluationcanada.ca/what-is-evaluation>
- Castonguay, A. (2016, November 28). Voici l'arme secrète de Trudeau. *L'actualité*. Retrieved from <http://lactualite.com/politique/2016/11/28/voici-larme-secrete-de-trudeau/>
- Cantelli, F., Jacob, S., Genard, J.-L., & de Visscher, C. (Eds.). (2006). *Les constructions de l'action publique*. Paris, France: L'Harmattan, Coll. Logiques politiques.
- Centre for Public Impact. (2016). *The public impact fundamentals report: Helping governments progress from idea to impact*. BCG Foundation. Retrieved from <https://publicimpact.blob.core.windows.net/production/2016/10/CPI-FUNDAMENTALS-REPORT.pdf>
- Centre for Public Impact. (2017). *A rubric for assessing public impact: Applying the public impact fundamentals to the real world*. BCG Foundation. Retrieved from <https://>

- publicimpact.blob.core.windows.net/production/2017/03/5388-CPI-Rubric-For-Assessing-Public-Impact-v3.pdf
- Centre for Public Impact. (n.d.). *What is public impact?* BCG Foundation. Retrieved from <https://www.centreforpublicimpact.org/what-is-public-impact/>
- CFA Local 1983. (2010a, April 20). *Why Deliverology made things worse in the UK. Part 2.* Speech by John Seddon at California Faculty Association Assembly [Video file]. Retrieved from <https://www.youtube.com/watch?v=6n9TH5ktKFE>
- CFA Local 1983. (2010b, April 20). *Why Deliverology made things worse in the UK. Part 3.* Speech by John Seddon at California Faculty Association Assembly [Video file]. Retrieved from <https://www.youtube.com/watch?v=2sIFvpRilSc>
- Chouinard, J. A., & Milley, P. (2015). From new public management to new political governance: Implications for evaluation. *Canadian Journal of Program Evaluation*, 30(1), 1–22. <https://doi.org/10.3138/cjpe.30.1.1>
- Couture, J., & Jacob, S. (Eds.). (2019). *Démocratie et politiques publiques*. Quebec, QC: Presses de l'Université Laval.
- Cummings, H., & Gauthier, B. (2016). *La SCÉ éduque l'Actualité à l'évaluation de programme*. Retrieved from <https://evaluationcanada.ca/fr/nouvelles/7306>
- Curran, R. (2016, April 27). Will deliverology work for the federal government? *Policy Options/Options politiques*. Retrieved from <http://policyoptions.irpp.org/magazines/april-2016/is-deliverology-right-for-canada/>
- Delacourt, S. (2017, November 21) *Stuck on hold: The AG's report and a government that can't get things done*. Retrieved from <https://ipolitics.ca/2017/11/21/stuck-hold-ags-report-government-cant-get-things-done>
- Delivery Associates. (n.d.a). *Who we are*. Retrieved from <https://www.deliveryassociates.com/who-we-are>
- Delivery Associates. (n.d.b). *Success delivered: How delivery units make a difference for governments and the citizens they serve*. Retrieved from [https://uploads-ssl.webflow.com/59ca37d5fcfb3000197aab3/5aa1526b45d91c0001b762b4\\_Inc%20London%20-%20DINO%20Brochure%20Amends%20-%20WEB.pdf](https://uploads-ssl.webflow.com/59ca37d5fcfb3000197aab3/5aa1526b45d91c0001b762b4_Inc%20London%20-%20DINO%20Brochure%20Amends%20-%20WEB.pdf)
- Dobell, R., & Zussman, D. (2018). Sunshine, scrutiny, and spending review in Canada, Trudeau to Trudeau: From program evaluation and policy to commitment and results. *Canadian Journal of Program Evaluation*, 32(3), 371–393. <https://doi.org/10.3138/cjpe.43184>
- Dyer, E. (2016a, January 19). Justin Trudeau's cabinet retreat stays on track until the end, almost. *CBC News*. Retrieved from <http://www.cbc.ca/news/politics/justin-trudeau-cabinet-retreat-new-brunswick-1.3411213> <http://www.cbc.ca/news/politics/justin-trudeau-cabinet-retreat-new-brunswick-1.3411213>
- Dyer, E. (2016b, April 26). "Deliverology" guru schools Trudeau government for 2nd time at cabinet retreat. *CBC News*. Retrieved from <http://www.cbc.ca/news/politics/deliverology-liberal-cabinet-retreat-1.3553024>
- The Economist. (2015, March 26). Improving government: Delivery man. *The Economist*. Retrieved from <https://www.economist.com/news/books-and-arts/21647263-tony-blairs-deliverology-expert-explains-how-reform-public-services-delivery-man>

- Fisher, C., & Downes, B. (2008). Performance measurement and metric manipulation in the public sector. *Journal of Business Ethics: A European Review*, 17(3), 245–258. <https://doi.org/10.1111/j.1467-8608.2008.00534.x>
- French, R. D. (2016, July 20). Trudeau's challenge: Turning promises into policy. *The Globe and Mail* (Ontario edition), p. A13. Retrieved from <https://www.theglobeandmail.com/opinion/trudeaus-challenge-turning-promises-into-policy/article30829997/>
- Geddes, J. (2016, October 31). Why McKinsey & Co. is helping Ottawa out, pro bono. *Maclean's*. Retrieved from <http://www.macleans.ca/politics/ottawa/why-mckinsey-helping-ottawa-pro-bono/>
- Global Agenda Council on Education. (2016). *Unleashing greatness: Nine plays to spark innovation in education* [White paper]. Retrieved from [http://www3.weforum.org/docs/WEF\\_WP\\_GAC\\_Education\\_Unleashing\\_Greatness.pdf](http://www3.weforum.org/docs/WEF_WP_GAC_Education_Unleashing_Greatness.pdf)
- Haws, E. (2018, January 24). Oft critiqued Liberal mandate tracker to become “more useful,” says government “deliverology” chief. *The Hill Times*. Retrieved from <https://www.hilltimes.com/2018/01/24/liberals-mandate-letter-tracker-great-conversation-getting-refresh-momentarily-mendelsohn/131770>
- Hood, C. (2006). Gaming in targetworld: The targets approach to managing British public services. *Public Administration Review*, 66(4), 515–521. <https://doi.org/10.1111/j.1540-6210.2006.00612.x>
- Hood, C. (2010). *The blame game: Spin, bureaucracy, and self-preservation in government*. Princeton, NJ: Princeton University Press.
- Hood, C. (2012). Public management by numbers as a performance-enhancing drug: Two hypotheses. *Public Administration Review*, 72(1), S85–S92. <https://doi.org/10.1111/j.1540-6210.2012.02634.x>
- Hood, C., & Dixon, R. (2015a). What we have to show for 30 years of new public management: Higher costs, more complaints. *Governance*, 28(3), 265–267. <https://doi.org/10.1111/gove.12150>
- Hood, C., & Dixon, R. (2015b). *A government that worked better and cost less? Evaluating three decades of reform and change in UK Central Government*. Oxford: Oxford University Press.
- Institute on Governance. (2018). *Results and performance bundle*. Retrieved from <https://web.archive.org/web/20190618163103/https://iog.ca/leadership-learning/course-bundles/results-and-performance-bundle>
- Jacob, S. (2006). Trente ans d'évaluation de programme au Canada : l'institutionnalisation interne en quête de qualité. *Revue française d'administration publique*, 119(3), 515–532. <https://doi.org/10.3917/rfap.119.0515>
- Jacob, S. (2009a). Opération chloroforme ou la réinvention de l'État rationnel : l'évaluation et les données probantes. *Criminologie*, 42(1), 201–223. <https://doi.org/10.7202/029813ar>
- Jacob, S. (2009b). Les systèmes de gestion de la performance et les mutations de la gestion publique contemporaine : enjeux et défis de la logique du classement. *Éthique publique : Revue internationale d'éthique sociétale et gouvernementale*, 11(1), 140–147. <https://doi.org/10.4000/ethiquepublique.1346>

- Jacob, S., Birch, L., Pétry, F., & Baby-Bouchard, A. (2018). Le gouvernement Couillard en tant que “mandataire indépendant.” In F. Pétry & L. Birch (Eds.), *Le gouvernement Couillard: 158 promesses et un mandat contrasté* (pp. 77–87). Quebec, QC: Presses de l’Université Laval.
- Jacob, S., Miller-Pelletier, A., & Birch, L. (2017). Les enjeux éthiques de la gestion par les résultats : les enseignements de la loi américaine No Child Left Behind. In B. Mazoux (Ed.), *Gestion par résultats : Concepts et pratiques de gestion de la performance des organisations de l’État* (pp. 173–192). Quebec, QC: Les Presses de l’Université du Québec.
- Jakobsen, M. L., Baekgaard, M., Moynihan, D. P., & van Loon, N. (2018). Making sense of performance regimes: Rebalancing external accountability and internal learning. *Perspectives on Public Management and Governance*, 1(2), 127–141. <https://doi.org/10.1093/ppmgov/gvx001>
- Lahey, R. (2017). Connecting the dots between M, E, RBM ... and deliverology. *Canadian Government Executive*, 23(1), 6–8. Retrieved from [http://canadiangovernmentexecutive.ca/dig/23\\_01/files/assets/basic-html/page-8.html](http://canadiangovernmentexecutive.ca/dig/23_01/files/assets/basic-html/page-8.html)
- Lahey, R., & Nielsen, S. B. (2013). Rethinking the relationship among monitoring, evaluation, and results-based management: Observations from Canada. In S. B. Nielsen & D. E. K. Hunter (Eds.), *Performance Management and Evaluation. New Directions for Evaluation*, 137, 45–56. <https://doi.org/10.1002/ev.20045>
- Laihonen, H., & Mäntylä, S. (2017). Principles of performance dialogue in public administration. *International Journal of Public Sector Management*, 30(5), 414–428. <https://doi.org/10.1108/ijpsm-09-2016-0149>
- MacCharles, T. (2017, November 14). Trudeau government’s mandate tracker lets you see if federal ministers are keeping commitments. *The Star*. Retrieved from <https://www.thestar.com/news/canada/2017/11/14/trudeau-governments-mandate-tracker-lets-you-see-if-federal-ministers-are-keeping-commitments.html>
- Manning, N. (2015). Supporting “small” improvements in government functioning: A rough guide for development professionals. In *A governance practitioner’s notebook: Alternative ideas and approaches* (pp. 151–189). Paris, France: OECD. Retrieved from <https://www.oecd.org/dac/accountable-effective-institutions/Governance%20Notebook%202.5%20Manning%20Small.pdf>
- Manning, N., & Watkins, J. (2013). *Targeting results, diagnosing the means: Innovative approaches for improving public sector delivery*. Recently Asked Questions Series. World Bank. Retrieved from <http://documents.worldbank.org/curated/en/136171467986259116/pdf/98815-BRI-VC-ADD-SERIES-PUBLIC-Box393182B.pdf>
- Mansbridge, J. (2003). Rethinking representation. *American Political Science Review*, 97(4), 515–528. <https://doi.org/10.1017/s0003055403000856>
- McGregor, G. (2016, May 24). Liberal government paid nearly \$18,000 for “deliverology” guru. *CTV News*. Retrieved from <https://www.ctvnews.ca/politics/glen-mcgregor-under-embargo/liberal-government-paid-nearly-18-000-for-deliverology-guru-1.2914536>
- Montague, S. (2016). *A response to does “deliverology” matter?: Targets can work if done properly*. Retrieved from <https://www.pmn.net/wp-content/uploads/Deliverology-Response-June-2016.pdf>

- Moynihan, D. P. (2005). Goal-based learning and the future of performance management. *Public Administration Review*, 65(2), 203–216. <https://doi.org/10.1111/j.1540-6210.2005.00445.x>
- Moynihan, D. P. (2008a). *The dynamics of performance management: Constructing information and reform*. Washington, DC: Georgetown University Press.
- Moynihan, D. P. (2008b). Advocacy and learning: An interactive-dialogue approach to performance information use. In W. Van Dooren & S. Van de Wall (Eds.), *Performance information in the public sector: How it gets used* (pp. 24–41). London, England: Palgrave Macmillan.
- Moynihan, D. P. (2010). A workforce of cynics? The effects of contemporary reforms on public service motivation. *International Public Management Journal*, 13(1), 24–34. <https://doi.org/10.1080/10967490903547167>
- Murekeyisoni, A. (2017). *La résultologie (Deliverology) comme instrument d'action publique: un cas ontarien* [Master's thesis]. École d'études politiques, Université d'Ottawa, Ottawa, Ontario. Retrieved from <https://ruor.uottawa.ca/handle/10393/36086>
- Nilsen, P., Ståhl, C., Roback, K., & Cairney, P. (2013). Never the twain shall meet? A comparison of implementation science and policy implementation research. *Implementation Science*, 8(1), 63. <https://doi.org/10.1186/1748-5908-8-63>
- OECD. (2010). *Quality standards for development evaluation*. Paris, France: OECD.
- Osborne, D., & Gaebler, T. (1993). *Reinventing government: How the entrepreneurial spirit is transforming the public sector*. New York, NY: Addison-Wesley.
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: SAGE.
- Pétry, F. (2014). A tale of two perspectives: Election promises and government actions in Canada. In E. Gidengil and H. Bastedo (Eds.), *Canadian democracy from the bottom up* (pp. 231–252). Vancouver, BC: UBC Press.
- Pétry, F., & Birch, L. (2016). Justin Trudeau's promises, one year later. *Policy Options/Options politiques*. Retrieved from <http://policyoptions.irpp.org/magazines/october-2016/justin-trudeaus-promises-one-year-later/>
- Pétry, F., & Birch, L. (Eds.). (2018). *Le gouvernement Couillard: 158 promesses et un mandat contrasté*. Quebec, QC: Presses de l'Université Laval.
- Pétry, F., & Duval, D. (2015). To what extent do political parties follow their campaign promises? *Journal of Parliamentary and Political Law*. Édition spéciale, 341–364.
- Poister, T. (2003). *Measuring performance in public and nonprofit organizations*. San Francisco, CA: Jossey-Bass, John Wiley & Sons.
- Puttick, R., & Ludlow, J. (2013). *Standards of evidence: An approach that balances the need for evidence with innovation*. NESTA (National Endowment for Science, Technology and the Arts). Retrieved from [https://media.nesta.org.uk/documents/standards\\_of\\_evidence.pdf](https://media.nesta.org.uk/documents/standards_of_evidence.pdf)
- Radwanski, A. (2017, November 17). Trudeau's "deliverology" on the verge of becoming a punchline. *The Globe and Mail*. Retrieved from <https://www.theglobeandmail.com/opinion/trudeaus-deliverology-on-the-verge-of-becoming-a-punchline/article37023338/>



- Richards, G. (2018). *Does deliverology deliver?* Unpublished Paper. Teffler School of Management, University of Ottawa.
- Richards, G., Gallo, C., & Kronick, M. (2017, January 30). *Does “deliverology” deliver?* Retrieved from <https://canadiangovernmentexecutive.ca/does-deliverology-deliver/>
- Ryckewaert, L. (2016, June 6). Privy Council’s new delivery unit increase capacity for centralized control, say experts. *The Hill Times*. Retrieved from <https://www.hilltimes.com/2016/06/06/new-pco-delivery-unit-increases-capacity-for-centralized-control-say-experts/67186>
- Sattler, P. (2012). Educational governance reforms in Ontario: Neoliberalism in context. *Canadian Journal of Educational Administration and Policy*, 128, 1–28.
- Savoie, D. (2015). *What is government good at? A Canadian answer*. Montreal, QC: McGill-Queen’s University Press.
- Schacter, M. (2002a). *Not a “tool kit”; Practitioner’s Guide to measuring the performance of public programs*. Retrieved from <https://www.files.ethz.ch/isn/122186/Guide.pdf>
- Schacter, M. (2002b). *What will be, will be: The challenge of applying results-based thinking to policy*. Retrieved from <https://iog.ca/research-publications/publications/what-will-be-will-be-the-challenge-of-applying-results-based-thinking-to-policy/>
- Schacter, M. (2006). *Trashing results-based management or throwing out the baby with the bath water*. Retrieved from [https://docs.wixstatic.com/ugd/dadb01\\_1a72e3b10b1e968a8943337920ad8e64.pdf](https://docs.wixstatic.com/ugd/dadb01_1a72e3b10b1e968a8943337920ad8e64.pdf)
- Schacter, M. (2011). *Tell me what I need to know: A practical guide to program evaluation for public servants*. Retrieved from [https://docs.wixstatic.com/ugd/dadb01\\_562bd8d94a9b0ca329730d21389d9b3e.pdf](https://docs.wixstatic.com/ugd/dadb01_562bd8d94a9b0ca329730d21389d9b3e.pdf)
- Schacter, M. (2016). *Does “deliverology” matter?* Retrieved from [https://docs.wixstatic.com/ugd/dadb01\\_67929f50ebaa46efa6e47df3526904a5.pdf](https://docs.wixstatic.com/ugd/dadb01_67929f50ebaa46efa6e47df3526904a5.pdf)
- Seddon, J. (2008). *Systems thinking in the public sector: The failure of the reform regime ... and a manifesto for a better way*. Axminster, England: Triarchy Press.
- Seddon, J. (2013). *Freedom from command and control: Rethinking management for a lean service*. New York, NY: Productivity Press.
- Shepherd, R. P. (2018). Expenditure reviews and the federal experience: Program evaluation and its contribution to assurance provision. *Canadian Journal of Program Evaluation*, 32(3), 347–370. <https://doi.org/10.3138/cjpe.43180>
- Smith, M. J., Richards, D., Geddes, A., & Mathers, H. (2011). Analysing policy delivery in the United Kingdom: The case of street crime and anti-social behaviour. *Public Administration*, 89(3), 975–1000. <https://doi.org/10.1111/j.1467-9299.2010.01892.x>
- Smithers, A. (2001). Education policy. In Anthony Seldon (Ed.), *The Blair effect* (pp 405–426). London, England: Little Brown. Retrieved from <https://www.buckingham.ac.uk/wp-content/uploads/2010/10/blaireduc.pdf>
- Tweedie, J. (2018, June 5). Does “deliverology” deliver the goods? *Canadian Government Executive*. Retrieved from <https://canadiangovernmentexecutive.ca/does-deliverology-deliver-the-goods/>
- Wells, P. (2016a, February 18). Meet Sir Michael Barber, the political delivery man. *Maclean’s*. Retrieved from <http://www.macleans.ca/politics/ottawa/meet-sir-michael-barber-the-political-delivery>

- Wells, P. (2016b, August 29). Cabinet committee changes give clues about happenings inside the Trudeau government. *The Hill Times*. Retrieved from <https://www.hilltimes.com/2016/08/29/cabinet-committee-changes-give-clues-about-happenings-inside-trudeau-government/78576>
- Welsh, D. T., & Ordóñez, L. D. (2014). The dark side of consecutive high performance goals: Linking goal setting, depletion, and unethical behavior. *Organizational Behavior and Human Decision Processes*, 123(2), 79–89. <https://doi.org/10.1016/j.obhdp.2013.07.006>
- Werner, A. (2018). Voters' preferences for party representation: Promise-keeping, responsiveness to public opinion or enacting the common good. *International Political Science Review*, 1–16. <https://doi.org/10.1177/0192512118787430>
- Wherry, A. (2016, August 27). How Justin Trudeau plans to deliver on “deliverology.” *CBC News*. Retrieved from <http://www.cbc.ca/news/politics/wherry-trudeau-deliverology-1.3735890>
- Wherry, A. (2017, November 14). Trudeau's Liberals check their to-do list: 67 promises down, more than 200 still to go. *CBC News*. Retrieved from <http://www.cbc.ca/news/politics/liberals-mandate-promises-deliverology-analysis-wherry-1.4400036>
- Zilio, M. (2016, September 30, updated May 18, 2018). British “deliverology” expertise costing Liberals \$200,000. *The Globe and Mail*. Retrieved from <https://www.theglobeandmail.com/news/politics/liberals-spend-200000-for-advice-on-delivering-campaign-pledges/article32187629/>

## AUTHOR INFORMATION

**Lisa Birch** is executive director of the Center for Public Policy Analysis, associate professor in the Department of Political Science at Université Laval, and professor at Champlain St. Lawrence College (Quebec City). Her expertise focuses on public management and public policy, textual analysis, and the fulfillment of campaign promises, including the methodology of the Polimetre, of which she is a co-founder.

**Steve Jacob** is a full professor in the Department of Political Science at Université Laval and director of the Center for Public Policy Analysis and PerfEval, a research laboratory on public policy performance and evaluation. Trained as a political analyst and historian, he conducts research dealing with the mechanisms of performance management and evaluation: professionalization, institutionalization and capacity building in Europe and Canada, and ethics in evaluation and participatory approaches. The results of his research have been published in numerous journals.