Sunshine, Scrutiny, and Spending Review in Canada, Trudeau to Trudeau: From Program Evaluation and Policy to Commitment and Results

Rod Dobell
University of Victoria

David Zussman
University of Ottawa

Abstract: *This review surveys experience with evaluation practices in the government of Canada since the mid-1960s, particularly with respect to spending reviews, concluding that there is little reason to expect any direct link from ongoing evaluation practices to cabinet decisions. The renewed commitment to evidence-based decision-making announced by the new Liberal government is unlikely to change this conclusion. The introduction of deliverology as a support function centred in the Privy Council Office shifts attention from policy formation to implementation and program delivery, with important emphasis on innovation and adaptation. But the crucial challenge still rests in achieving greater public access to information and greater inclusiveness in decision processes. For academic leaders in public administration, attention now should shift from terminological and doctrinal disputes to anticipating the important consequences of machine learning and artificial intelligence for education and future professional practice in public policy.*

Keywords: *deliverology, spending reviews, public expenditure, access to information, open governance, artificial intelligence*

Résumé: *Cet article fait état du rôle de l'évaluation au Gouvernement du Canada depuis le milieu des années 1960. En particulier, il analyse le lien entre l'évaluation et les exercices de révision budgétaire, concluant qu'il y a peu de raisons d'attendre de l'évaluation qu'elle ait une influence directe sur les décisions du Cabinet. L'engagement renouvelé du nouveau gouvernement libéral à ce que les décisions soient fondées sur les données probantes ne changera probablement pas cette conclusion. L'introduction de la résultologie, comme une fonction en appui au Bureau du Conseil Privé, détourne l'attention de la conception des politiques vers l'implantation et l'exécution des programmes, en mettant l'accent sur l'innovation et l'adaptation. L'enjeu principal reste toutefois le même : réussir à améliorer l'accès public à l'information, et à être plus inclusif dans les processus décisionnels. Pour*

Corresponding author: Rod Dobell, University of Victoria, Centre for Global Studies;
rdobell@uvic.ca

les leaders académiques de l'administration publique, l'attention devrait maintenant passer des disputes terminologiques et doctrinales à l'anticipation des conséquences importantes de l'apprentissage automatique et de l'intelligence artificielle pour l'éducation et la future pratique professionnelle dans la politique publique.

Mots clés: *résultologie, revue des dépenses, dépenses publiques, accès à l'information, gouvernance ouverte, intelligence artificielle*

INTRODUCTION

The purpose of this article is to examine what the authors see as a distinct shift in context and emphasis from the early years of program evaluation, based on policy sciences, to current views of evidence-based decision-making, designed to serve governance in a postmodern, post-fact world. The paper does not purport to offer fresh data or original theory; its more modest goal is a brief expository survey. The question it confronts is therefore simply to know how we might view the evolution of evaluation and spending review in the government of Canada over the 50 years since the efforts of Prime Minister Pierre Trudeau to bring rational analysis into management in that Government.[1]

In 1969, driven by the "whiz kids" in the United States and the 1960s wave of enthusiasm for systems analysis and "modern" government, the then-new Canadian government of Prime Minister Pierre Elliot Trudeau adopted the Planning, Programming, Budgeting System (PPBS) to guide its expenditure management decisions. This work was to be led by the Treasury Board Secretariat (TBS), recently moved from the Department of Finance to be a separate department under its own minister, the president of Treasury Board. To support the introduction and implementation of this new thinking, a new unit, the Planning Branch, was created within the secretariat. Appointed as deputy secretary to head this new unit was Douglas Hartle, professor of political economy and director of the University of Toronto's Institute for Policy Analysis, and formerly the director of research for the massive Royal Commission on Taxation (the Carter Commission).

Now, almost 50 years on, driven by new reformers in the United Kingdom and the recent wave of enthusiasm for results and accountability, the government of Prime Minister Justin Trudeau[2] has adopted *deliverology* to ensure realization of the government's promises.[3] In its campaign platform and subsequent to the 2015 election, the new government committed itself to evidence-based decision-making both in formal cabinet-level decision processes and throughout government operations, including program delivery.

This work was to be guided by a new cabinet committee on agenda and results, chaired by the prime minister, together with a new deputy ministers' committee on policy innovation. To support the introduction and continuing implementation of this new thinking, a new unit, the Secretariat on Results and Delivery, was created within the Privy Council Office. Appointed as deputy secretary to head this new unit was Matthew Mendelsohn, director of the

University of Toronto's Mowat Centre for Policy Innovation, formerly deputy minister in the Ontario government and co-author of the 2015 Liberal Party election platform.

That platform committed the government to an expansive budgetary posture and substantial fiscal stimulus, leading to the prospect of large budget deficits sustained over several years. Budget 2017 offered no timeline for return to budget balance, though aiming at a declining debt/GDP ratio over the period to 2019. Since then, a new defence policy paper (June 6, 2017) suggests—perhaps in response to American pressure in NATO—very significant increases in defence expenditures. The budgetary cost of the government's commitment to reconciliation with First Nations and Indigenous people has yet to be reckoned.

There appears to be no appetite at this time for a dramatic expenditure reduction and cost control exercise, such as the 1994 Chrétien–Martin program review initiative that successfully addressed the massive budget deficit left by the previous Mulroney–Campbell government, or even the subsequent strategic reviews and Deficit Reduction Action Plans (DRAP) of the Harper government. Nevertheless, the possible need for such an initiative has to be anticipated. In that case, can either program evaluation, based on systems analysis, or the machinery of deliverology, really help?

We accept the general conclusions in the articles by John Mayne and others in this special issue that formal program evaluation cannot help much, and we argue below that deliverology may be a different thing. We will therefore not attempt any survey or appraisal of the current academic or professional literature on the theory and practice of evaluation—topics that are pursued more fully and effectively through other pages of this journal, as well as in reviews such as Dobell (2003) or in standard texts such as Eliadis et al. (2011) or McDavid, Huse, and Hawthorn (2012). Rather, we return to the message of an earlier paper (Dobell and Zussman, 1981) to emphasize that the key concerns lie elsewhere, in the role that must be played by disclosure of the narratives and evidence underlying decision-making (including performance monitoring and program evaluation results) in buttressing the legitimacy of government. For this reason, we emphasize the need for the present government to complete the crucial review and reform of access to information policy that was promised in the 2015 election campaign to offset the erosion of that policy that has occurred since it was developed in the earlier Trudeau government in the 1970s and 1980s.[4]

BACKGROUND

To make sense of all the recent activity and appraise its significance, we begin with some distinctions, separating several different topics. The first includes evaluation and challenge functions that deploy explicit—some might say elite or technocratic—analysis in support of executive decisions, either to initiate or to terminate programs or activities. These functions generally emphasize what is called summative evaluation, addressing go/no-go decisions for individual, large-scale

programs. Analytical support for these decisions might include ongoing program evaluation (with evolving criteria, as exemplified in the 2009 Treasury Board Secretariat assessment and reorientation [Canada, Treasury Board Secretariat, 2009] of evaluation activities that emphasized the seemingly more politically charged concern for "relevance" as a central criterion in [then] future evaluation efforts). Or they might offer a challenge function, where the considerations relate to the choice of organizational or institutional form for delivery of program outcomes, the test here being whether the activity under review should, given its ideological orientation, be undertaken by the government at all.

Within this framework, at cabinet level, one might expect to look to summative evaluation in strategic planning or in the ex-ante appraisal of proposed new programs, or to centre-led ex-post testing of the continued effectiveness, relevance, and value of existing ongoing programs, possibly in the context of massive program reviews and spending cuts.

Second, tracking—either internally or in public—of progress toward government fulfillment of election commitments and implementation of decisions by cabinet or departmental management should be seen as different, and is one aspect of deliverology. Of course, a range of institutional frames and incentive structures has long existed inside the federal government to promote tracking and delivery in programs approved for continuation, well before the deliverology label emerged (Curran, 2016).

A third category, aspiring to a regime of continuous improvement at a more operational level, spans monitoring and tracking of performance inside government in the delivery of programs. At this more operational level, within departments, there has long been an array of mechanisms for tracking ongoing management of programs within the public service. This grouping includes performance measurement and monitoring systems such as the canonical Operational Performance Measurement System promoted by the Efficiency Evaluation Division of the Treasury Board Secretariat Planning Branch in the early 1970s and used in the government of Canada in some form ever since.

To guide the evolution of continuing government programs, management can rely on a schedule of ongoing formative evaluation over its full range of activities. This continuous operational monitoring and appraisal function is distinct from the more fundamental periodic evaluations mentioned above (Maxwell, 1986; McDavid et al., 2012), and indeed could be seen as a second dimension of deliverology, tracking ongoing follow-through of cabinet decisions into continuing departmental operations as a new responsibility at cabinet committee level, with the support of the Privy Council Office, Results and Delivery unit).

As a fourth topic, a related development that also falls under the label of deliverology should be highlighted as a distinct feature of the new thinking in Ottawa. That is the emphasis on behavioural psychology and behavioural economics in the work of so-called nudge units and innovation hubs (Chen, Bendle, & Soman, 2017). These might be seen as contributing to formative program design and (perhaps quite subtle) adaptive management decisions. But at the scale of the federal

government there are challenging institutional barriers to the style of adaptive management implied by emphasis on this work (Longo, in press).

Fundamental to all of this, we argue, is a distinct category of activities related to accountability mechanisms going beyond the direct interests of the executive in tracking its own performance. It involves the never-ending concern with information for Parliament and its officers, general reporting to the public on government spending, and other information for public consumption,[5] as well as, now, in the other direction, the growing fascination with open data, open government, and monitorial citizenship (Graeff, 2017) as mechanisms for holding governments to account and establishing the legitimacy that flows from transparency and access.

And, finally, as will be mentioned briefly in our concluding comments, we have to look now not just to access by citizens to information, but to participation of citizens in decision processes, reflecting ideas of open democracy and collective intelligence enabled by the transformation of information and communications technologies (Landemore, 2013, 2016; Sevinc, 2017).

Despite this impressive array of machinery, of course, sceptics argue that it is crucial to ask where is the demand for such work and where might the information generated by all this activity actually be put to use? If neither public nor Parliament express any appetite for evidence-based decision in this post-fact world, why should we expect any of the apparatus described above to enter the real life of the public servant, let alone alter political decisions in ongoing operations or spending reviews? The recent exhaustive review by Rick French (French, 2018a,b) of actual experience in this regard should be required reading for anyone interested in this question.

HISTORICAL CONTEXT

In this section, we briefly review the high points in the evolution of practice—or at least intention—over the past 50 years, first with respect to ongoing evaluation practice and then looking to the particular case of intermittent comprehensive spending reviews, then deliverology, and, finally, developments toward greater transparency, with accounting officers, new oversight agencies, and the emerging emphasis on reporting and accountability as a distinct concern.

Program Evaluation

In the 1960s and 1970s, in the aftermath of World War II, enthusiasm for operations research and systems analysis, with the RAND Corporation appearing to solve intractable problems in economics and public sector management, Western governments enthusiastically embraced Planning, Programming and Budgeting Systems (PPBS), with its supporting analytical machinery, including cost-effectiveness or cost–benefit analysis, to pursue optimal policies, or at least to increase the effectiveness of resource allocation decisions.

In 1966, the Treasury Board Secretariat became a stand-alone department of the government of Canada, responsible for expenditure management, for moving

cabinet decisions into government operations, and for implementing the recommendations of the Royal Commission on Government Organization (the Glassco Commission), established in 1960 and reporting in 1963. The central message of that commission was to "let the managers manage" (in a rational manner), arguing that departments should be free from excessive central control and should be encouraged to devise management methods suited to their needs.

The government's response, as set out in the Treasury Board guide to financial management (Canada, Treasury Board, 1968), together with the two prior policy statements on performance measurement, in 1965, and financial management, in 1966, might be taken as the starting point for our brief survey here.

Reflecting the objective of professionalizing the public service by implementing modern management techniques, the Planning Branch of the Treasury Board Secretariat was formed in 1969 to build the analytical support for government decisions founded on "reason before passion." That branch was organized in three main divisions: efficiency evaluation (performance measurement and monitoring), effectiveness evaluation (program evaluation, including appraisal of ongoing programs and assessment of proposed initiatives), and organization (looking to appraisal of proposed structural or institutional innovations, as well as appointment of senior personnel).[6]

Management systems and policies on evaluation have since been through many stages in the government of Canada; the story is far too long to recount here. (See Segsworth [2005] for an account of events to 2005, Shepherd (2012), and Canada, Treasury Board Secretariat (2013–2014) for an exhaustive overall summary; Shepherd (2018) also covers policy changes over time.) With respect to prescribed evaluation practice, a benefit-cost manual was issued by Treasury Board in 1977, and in the same year a first set of policies and directives on evaluation (based on the framework described in Hartle [1973]) was formally set out. Over the subsequent 3 decades these went through a series of revisions and adjustments, then on to a serious reappraisal in 2009, reflected in the new *Policy on Results* introduced, with substantial fanfare and extensive training efforts, by the new government in July 2016.

Program evaluation has thus been treated, under the guidance (or supervision, depending on the point of view) of the Treasury Board Secretariat, as a mandatory part of ongoing activity in Ottawa since 1966. This action ushered in a new industry for consultants and academics, who flocked to Ottawa to be part of this exciting new activity that was designed to make government more rational, effective, and efficient.

Over the past 2 years, the current federal government has taken several steps intended to strengthen the program evaluation function. It has released a new *Policy on Results* (with a new but not novel directive on program evaluation) requiring departments to evaluate their programs on a 5-year cycle (see Canada, Treasury Board Secretariat, 2016a, 2016d); it has appointed chief evaluation officers in all federal departments, who report directly to the deputy minister and who have full responsibility for ensuring that comprehensive evaluation

plans are prepared and carried out[7]; and it has apparently embraced thoroughly the philosophy of deliverology as preached by Sir Michael Barber to Tony Blair (Zilio, 2016).

The seriousness of the government's intentions was underlined with the release of the prime minister's mandate letters to the president of the Treasury Board and the minister for democratic institutions. These letters explicitly assigned them responsibility to "work with the Minister of Finance and your colleagues to conduct a review of tax expenditures and other spending to reduce poorly targeted and inefficient measures, wasteful spending, and government initiatives that are ineffective or have outlived their purpose,"[8] and "take a leadership role to review policies to improve the use of evidence and data in program innovation and evaluation, more open data, and a more modern approach to comptrollership" (Canada, Office of the Prime Minister, 2015).

Beyond the federal government, it is important to note current work toward a broader conceptual and procedural framework that, while it may not be well-suited to deliberation at cabinet level, offers real promise at a sectoral or regional scale. Hodge (2017) illustrates the potential for path-breaking participatory contribution analysis using applications in the mining and minerals sector as examples; Budhwani and McDavid (2017) provide a contemporary survey of contribution analysis more generally.

Spending Reviews

Through the early 1970s, things were generally buoyant in Canada and abroad. The A/B/X approach to budgeting embedded in PPBS focused on B: discretionary new programs were embraced enthusiastically. The A-base of continuing programs expanded with inflation, demography, and increasing participation rates (see Canada, Treasury Board, 1978). And the X-budget, identifying expenditure reductions, received little attention, despite in some cases enthusiastically negative ratings flowing from summative evaluation efforts.

This all changed in the late 1970s, most visibly in 1978 when the prime minister returned from the G7 summit meeting in Bonn to inform his minister of finance and Treasury Board president that he had undertaken a commitment to a substantial budget cut. The system had to scramble to deliver.

With the election of Brian Mulroney in 1984, the Ministerial Task Force on Program Review led by Deputy Prime Minister Erik Nielsen followed, with the professional support of the public service and the input of external consultants who were engaged to take a whole-of-government look at annual spending. As a result, after 18 months of work, over 100 private sector participants and as many public servants reviewed more than 1,000 programs. While the analysis yielded many potential areas of government savings through the termination of existing ineffective programs and merging of overlapping activities, Prime Minister Mulroney did not have the heart to make the tough spending cuts and little financial saving was achieved despite the concerted efforts of many experts (Bourgon, 2009).

It wasn't until 1994, with the return of the Liberals under Prime Minister Chrétien, that another serious assessment of government spending was undertaken. This time, a sluggish Canadian economy and skyrocketing national debt forced the prime minister to take bold action. The resulting exercise, known as program review, was championed by a cabinet committee, chaired by Marcel Massé, president of the Treasury Board, but effectively driven by the public service, particularly by the Department of Finance, who established hard savings targets for each department and agency. As well, the committee established a requirement that every federal program be subject to an assessment based on six simple questions about their overall contribution to the public good (see Appendix 1).

Shortly after coming to power in 2006, the Harper government announced several strategic spending reviews. The purpose was to demonstrate that embedded in federal programs was a lot of wasteful or inefficient spending that could be diverted from underperforming programs to those that reflected the Conservative government's own priorities. All government departments and agencies were subjected to an ongoing cost-cutting exercise with the intention of achieving a 5% reduction in government spending. Five years later, the Harper government reported that 98% of direct program spending had been reviewed and that the 5% savings target had been achieved (Savoie, 2013, p. 102).

In 2011 the Harper government launched a second review designed to be more strategic and targeted toward identifying $5 billion in annual savings by 2014. This time, as compared to the Neilsen task force, which brought in seconded (volunteer) private sector personnel to work with the public servant teams, the government contracted the whole cost reduction exercise to a consulting company. The idea was to remove public servants, perceived to be self-interested, from the review and to encourage government-wide recommendations from outside advisors thought to bring a more independent and comprehensive view. At the political level, the DRAP exercise was led by a special Treasury Board cabinet committee that also included a large number of experts from outside of government who participated in the decision making. When confronted about the appointment of people from outside the government to lead an internal cabinet task force, Finance Minister Jim Flaherty justified the $20 million contract with Deloitte Touche by saying that "it isn't good, quite frankly, for government to look just at itself" (Curry, 2011).

Deliverology (Delivery and Results)

Moving in a different direction, a central innovation hub was promised in the 2013 federal public service review initiative that led to the *Blueprint 2020* report (Canada, Privy Council Office, 2016). In 2014, one such central innovation hub was opened in the Privy Council Office, with hubs also in Employment, Social Development and the Canada Revenue Agency to pursue the model of the UK nudge units focused on experimentation and the co-design or collaborative tweaking of program designs (Chen et al., 2017). This focus on experimentation introduces a new enthusiasm to be pressed on the public service: in place of the

marketing, early in the Trudeau I era, of all the esoteric machinery of cost–benefit analysis in all its variants—including the risk–benefit analyses on which it usually founders—we have now, during Trudeau II, all the esoteric machinery of the "gold standard" of randomized controlled trials (or the ingenious mutations envisaged when the strict conditions for RCT are missing) as the subject of the uphill missionary work radiating out from the centre (Canada, Treasury Board Secretariat, 2016b), this time from the Privy Council Office rather than the Treasury Board Secretariat.

A rationale for the "Libertarian Paternalism" underlying this work was set out by Thaler and Sunstein (2008), and found its way also into the Obama White House (United States, National Science and Technology Council, 2017), though not, perhaps, to the same extent as in 10 Downing Street. The basic idea is that one can affect behaviour by shaping choice rather than regulating action. As a result, coercive regulations and the threat of enforcement action for violation of them are replaced by a reframing of the terms of the choice to be made, nudging and leaning to recast the way one sees one's own preferences and thus to shape consequent voluntary action. The spreading impact of this idea is reflected in the award of the 2017 Nobel Prize in Economic Sciences to Richard Thaler (and, before that, to Daniel Kahneman in 2002).

As noted above, it is important to distinguish this work on experimentation, nudging, and adaptive management from the tracking of results and delivery that is the other face of deliverology, and more closely related to the next topic.

Reporting and Accountability

There has been overarching change in public pressure and expectations for accountability and transparency in government. The first years of the Trudeau I government saw the end of the substantially untroubled post–World War II reconstruction and expansion in Western economies and the beginning of the period that led the OECD to launch the McCracken committee, its international enquiry into "What went wrong?" in the global economy. In Canada, concerns about rising deficits led to anxiety in the business community.

This concern coincided with growing disputes between the government and the auditor general—first Maxwell Henderson, then James J. MacDonnell—over the right of the auditor general to comment on the wisdom of government decisions. A committee, the Independent Review Committee on the Office of the Auditor General (the Wilson committee), was formed. It recommended a new Auditor General Act to extend substantially the scope and powers of the auditor general, in particular, creating a mandate to go beyond traditional audits, attesting that the books of the government fairly represented its financial position and that monies had been expended for the purposes for which they had been appropriated by Parliament.

The notion of "management audit" was introduced, inviting the auditor general to report on whether "value for money" had been received as a result of the decisions of the government. Under the new act, James MacDonnell took those

new powers seriously. Beyond his famous remark—"I am deeply concerned that Parliament—and indeed the government—has lost, or is close to losing, effective control of the public purse"—MacDonnell pressed heavily to move the auditor general role deeply into what had always been considered appraisal addressed to political, not accounting or technical, questions. The accountability of the government on economic and social, as well as financial, issues became a controversial topic.

Accountability on other fronts also became topical with the Royal Commission on Financial Management and Accountability (the Lambert Commission), which issued its final report in 1979. The memorable tag line from that report was "make the managers manage," suggesting that decentralization and delegation in the federal government, as recommended by the Glassco Commission, had gone too far, and greater oversight was essential.

The 1982 report of the Special Parliamentary Task Force on Fiscal Federalism similarly highlighted the concept of accountability as crucial in deciding on the appropriate division of responsibilities between national and subnational governments.[9]

This pressure led in 2006 to the creation by the incoming Harper government, as part of the Federal Accountability Act, of the Parliamentary Budget Office, the public sector integrity commissioner, commissioner of lobbying, and the conflict of interest and ethics commissioner, as well as the formation of departmental audit committees (Zussman, 2017). At the same time, there was an increase in the oversight capacity of several parliamentary agents and a growing network of independent think tanks and other types of organizations dedicated to tracking government activities, providing parliamentary oversight (watchdog roles) and policy commentary.

Over the past decade, the pressure to move from inside scrutiny to outside has continued even more strongly. A flow from confidence in program evaluation to reliance on disclosure of information can be traced strikingly in the work of Doug Hartle from the first days of the Planning Branch to his later advocacy back at the University of Toronto (Dobell, 1999), and in subsequent literature leading recently to notions of monitorial citizenship (Graeff, 2017) and expansion of independent web-based "tracking" services (e.g., Promise Tracker [n.d.] or TrudeauMeter [n.d.]).

DISCUSSION

Over the past 40 years or so there has been a steady stream of research and commentary on the state of evaluation at the federal level. In 1981, Dobell and Zussman reviewed the progress and contribution of efforts at evaluation activities over the prior decade (Dobell and Zussman, 1981), noting that there seemed little evidence of success as measured against promises and expectations. By 1986, the Nielsen task force concluded that "government program evaluations were generally useless and inadequate despite the emergence of evaluation units in most government departments" (Canada, Task Force on Program Review, 1986, p 23).

The piling on continued in the academic literature when Peter Aucoin observed that program evaluation "has never been embedded in the expenditure management decision-making process" (Aucoin, 2005, pp. 5, 10), or with Bourgeois and Whynot (this issue), who concluded that "our findings are generally consistent with those of previous studies and reveal that, although instrumental and conceptual use tends to be quite high at program level, there is little or no strategic use of evaluation beyond program improvement, and that evaluation is not used to inform spending reviews" (2018, p. 342). Perhaps the debate can be closed off with a summary judgement offered by Barry Carin, both a major supplier and *demandeur* of evaluation work and a founder of the Canadian Evaluation Society, which publishes this journal:

> On balance, I conclude that it is an illusion to think that you can be analytic and positive—there are too many incommensurable objectives with incomplete metrics, too many permutations and combinations of pertinent sensitivity calculations beyond what anyone can absorb, let alone have the patience to read. I therefore do not believe that you can either perform evaluations useful for Ministers, or evaluate the evaluation function (as I just did). (B. Carin, personal communication, June 2017)

Thus, with respect to the direct influence of program evaluation efforts on routine decisions, the lesson is generally accepted, as John Mayne (2018) perhaps confirms. It now seems necessary to fall back on the larger vision, the article of faith driving commitment to evaluation in government from the 1960s onward. This faith is that ongoing systematic evidence-based analysis builds a foundation of knowledge and enlightenment (Weiss, 1972, 1999) from which spring the relevant stories, beliefs, and perceptions on which intuitive decisions will be based.[10]

What about spending reviews themselves? Is there a case for any useful link from program evaluation to central spending reviews in situations of critical budgetary pressure? It seems not. The evaluation work of the Planning Branch played no role in the 1978 X-budget. Indeed the branch was just in the process of being dissolved. "The 1986 Nielsen Task Force was effective in reducing expenditures where rational evaluation aligned with the quiet suspicions of senior managements" (H. Swain, personal communication, June 2017). This observation nicely reconciles the comment quoted above, that task force study teams found evaluations "useless," with the protest by John Mayne (Mayne, 1986) that evaluations had been extensively used.

While the 1995 program review budget specifically targeted underperforming programs, there is no evidence that Prime Minister Chrétien, Finance Minister Paul Martin, or Treasury Board President Marcel Massé used program evaluation reports in identifying programs for cancellation or spending reductions. But is all this surprising? In routine circumstances, something like the old A/B/X approach to budgeting can prevail: the ongoing operational A-base moves along incrementally, the B-budget to fund new programs is allocated, and the X-budget identifies programs that can be eliminated to make room for the new.

In times of crisis, however, it is essential to go beyond the decremental or marginal X-budget, and dig deep into the A-base to regain control of total expenditures while funding the new activity essential to a growing, changing, and innovative society. In the compressed time frame of such a spending review, there can be no realistic expectation of directly relevant evaluation results on tap or built to order that can bring information to bear on judgments about the relative importance or relative priority across broad categories of program expenditures. The best that one might hope is that triangulation across the perceptions, beliefs, and views—with luck, perhaps influenced by a foundation of accumulated past individual evaluation efforts—of diverse groups of responsible ministers can provide a basis for consensus on action-oriented decisions.

There is, of course, an even broader tension underlying all these efforts to convey information on which to ground action. The biggest challenge, as always, is flawed implementation of decisions and intentions—particularly where those intentions are vast and "aspirational." Implementation of policy intentions demands performance by individual public servants exercising discretion within a body of public service norms. It demands fidelity to the text expressing those intentions, but it also demands interpretation of that text and exercise of discretion in carrying out the performance.

The problem—an ancient problem—is whether to exercise that discretion in accord with the black letter text or according to the dictates of particular changing and uncertain circumstances, on the ground. This is where the fundamental challenges of accountability arise, but also where the constraints of excessive audit bind badly. The search for the "sweet spot" between "letting the managers manage" according to their instincts, within general understandings of the rules, and "making the managers manage" within clear and precise uniform rules established in light of consistent intentions across the enterprise seems never satisfactorily ended.[11] A wonderful example is provided by the controversy around the public service attempt in 2017 to codify the operational implications of a 2011 tax court ruling on the status of employee discounts as a taxable benefit for income tax purposes (Tasker, 2017).

Can a new analytical unit add anything significant to the delivery-oriented monitoring structure already inherent in the institutional structure of the government of Canada? Are there better ways to decentralize responsibility and move monitoring from the centre? It can be argued that deliverology, as experimentation, innovation, and adaptive management, while promising in principle, is not well adapted to the role of the federal government, which has control of very little direct delivery of services (Curran, 2016)—and where it does, as in the Phoenix payroll system, it does not have an impressive record.

The really significant shift does not rest in the new *Policy on Results*, still based on the idea of the rational economic man, but rather in a deeper intellectual transition recognizing the imperfections in human reasoning, reflected in the rise of behavioural insight units, and behavioural economics more broadly. We return briefly to this question below.

In the first decade of the Pierre Trudeau administration, the focus—at least for the analysts and "hot-shot advisors"—was on the policy formation phases of the policy process. Once cabinet had made—or not—a decision, that was the end of the story for analytical purposes (though not, of course, for the vast majority of public servants, whose task was simply to keep the turbines of the power company turning while those decisions were implemented). The work to be done in implementation—realizing the intent of decisions endorsed—was taken for granted as routine management (even though in the middle of that decade Pressman and Wildavsky [1973] published what was to become a famous book on implementation with a most suggestive subtitle, *How Great Expectations in Washington Are Dashed in Oakland*).

Now in the first term of the Justin Trudeau government, the focus seems to have shifted significantly to the implementation and delivery phase of the cycle. This shift matters beyond the peculiar interests of the academic observer; for the evaluation industry and evaluation practice, it moves the emphasis away from summative evaluation and decisions, and toward policy design, formative evaluation, and adaptive management (Gold, 2014). For the public service it forces attention to the changes in policy that actually occur at the workface, at the level of individual programs, and, indeed, at the level of design of forms in individual sub-activities.[12] And this is where behavioural insights come into play.

CONCLUSION

We have argued that the current structure for evidence-based management by results in the federal government is, despite some changes in labels and story lines, not very different in purpose or substance from what has been familiar for a long time. With respect to evaluation policy and evaluation activities generally, we do not dispute the general opinion expressed elsewhere in this volume that there is no reason to expect any direct link from program evaluation to cabinet decisions. While evaluation efforts addressing ongoing programs—formative evaluation— can offer obvious contributions to continuous improvement in departmental expenditure management, and nudge units can suggest fruitful reframing of designs, they are highly unlikely to offer empirical or analytical evidence that ministers will see as definitive or even terribly relevant to specific cabinet decisions. What can be hoped, as emphasized by one of the founders of the program evaluation movement in the United States, Carol Weiss, is that a culture of ongoing analysis and evaluation will build a foundation of general understanding and "enlightenment" that informs ministerial perspectives and cabinet deliberations (Weiss, 1972, 1998).

We have argued more specifically, on the basis of extensive personal experience, that in the case of either occasional or systematic spending reviews, neither the timetable nor the political context lends itself to considered comprehensive reduction of the expenditure budget as a package. Decisions about cuts will likely be made arbitrarily, based on perceptions and beliefs—probably those of one strong

individual enjoying the confidence and backing of the prime minister—about program merits but, more importantly, about the appropriateness of the activity as a responsibility of the federal government of the day. Again, perhaps the most one can hope (counterfactually?) is that an ongoing culture of respect for evidence and interest in outcomes will have shaped the opinions and perspectives of that central individual in a way that promotes a broad public interest rather than ideologically-driven motivated reasoning pressing toward ever-smaller government.

This leaves us with hopes, first, for effective tracking of commitments and monitoring of progress internally within departments and government as a whole, and visible public accountability for continuing and faithful implementation of decisions and realization of commitments. Much of this machinery has long been in place in the federal government, but perhaps it is here that deliverology (under whatever label) and the cabinet committee on agenda and results will concentrate the minds of politicians and managers and thus offer fresh impetus to the concern for effectiveness in the delivery of just outcomes.[13] Beyond that, there is the hope that effective measures for reporting to Parliament will generate feedback that will forcefully shape ministerial and managerial perceptions as to the public interest. Those attitudes might thus be shaped indirectly through program evaluation and the culture of attention to the evidence underlying stories and narratives.

None of this can really work, however, without a dramatic transformation in current government attitudes toward public access to information in government hands and to the use being made of that information in shaping government decisions. Despite all the protestations of openness and transparency, practice in this matter is disappointing. The recent experience of the auditor general in being denied access to analysis of tax expenditures associated with the government's commitment to phase out "inefficient" subsidies to fossil fuel industries, and indeed failure to receive documentation on action taken toward implementation of that commitment, is only one extreme example.[14] As we argued in 1981, we need sunshine and scrutiny, and we need access for the public to see the evidence and the analysis on which government decisions are built and on which the legitimacy of government must be maintained.

But of course, this is not 1981. Government attitudes toward openness now must reflect the massive transformations since then in technology, information and communications systems, social networking, learning algorithms, and artificial intelligence (see Longo, in press). Can this Liberal government succeed in translating the basic principles of the access to information policy introduced by the Liberal government of the 1980s into an appropriate information regime for this new context? If so, government's role in evaluating government (for decision purposes) may prove much less significant than the public's role in evaluating government (for accountability and legitimacy) on the basis of fuller disclosure. "It is possible . . . that institutional arrangements to strengthen representative institutions could be developed in the context of a participatory e-democracy. In such a setting, a deliberative democratic approach to evaluation could be pursued

in inclusive processes at various scales, as part of ongoing conversations, ongoing social learning, ongoing adaptive management, in ongoing resilient communities" (Dobell, 2003).

Because it is 2017, we need now to frame the case for access much more expansively. In an era of Big Data and open government, we must argue passionately for open public access not just to the data assembled and held by governments but also to the interpretations of those data buried in the stories and narratives built upon them for purposes of providing background to government decisions. This access need not reach to the options identified or recommendations offered by public servants in memoranda to cabinet, or to opinions voiced in cabinet deliberations, all of which rightly must remain cabinet confidences. But it should include the analysis and interpretations built by public servants from data in public hands. And, indeed, beyond access to these materials, arguments for opportunity to participate in the processes of framing the questions themselves and of articulating the context for creating these stories and narratives must be vigourously pursued (see Chouinard, 2013; Chouinard & Milley, 2015)

More demandingly still, not only must the case for access to the data and information in government hands, and to interpretations of them—the evidence, representations, and visualizations on which decisions are built—and to the process of framing those questions and creating those interpretations, be maintained, but also, finally, the case must be made for inclusive citizen participation in the process of building legitimate collective decisions themselves from those interpretations and stories.[15]

And, to close, yet more speculatively, we as academics or practitioners should perhaps ask ourselves how long we should expect to go on as we are in our promotion and practice of current theories or techniques of evaluation. It is a new world out there; maybe it should also become a new world in here, in the academy and the profession. How long might it be before the whole process of extracting relevant evidence from vast banks of observations and data will be handled better by smart robots than by humans? Given the extent to which audit, accounting, analysis, evaluation, and performance measurement are generally taught through established recipes and simple inference or learning processes, surely it is likely that recursively self-improving automata will quickly learn the required skills? That this outcome may arrive sooner than we think is suggested by evidence already with us.[16]

Obviously, these further ideas take us into a whole new world, well beyond the scope of the present note. But perhaps in closing we can once again urge the evaluation community away from continuing to differentiate the labels and fine-tune the footnotes or nuances in a literature on the main conclusions of which—as summarized above—academicians and practitioners generally agree. Rather, perhaps we might turn to exploring, tentatively but vigorously, the vast new worlds that open to us to transform our representative democracy by engaging the people in its decisions.

NOTES

1 The authors have between them many years of academic research and hands-on experience in or around this field, particularly at the federal level.

2 The excessive personalization of the Government of Canada embedded in this shorthand is regrettable, but seems necessary to avoid cumbersome circumlocution.

3 Distinct components of the machinery that goes under the label 'Deliverology' are discussed below.

4 After this article was largely completed, we received the annual report of the information commissioner for 2016–2017. It seems essential to highlight the message with which she leads off that report:

> In March 2017, the government announced its plans to delay the first phase of the Act's reform, citing the need to "get it right." Our investigations reveal, once again, that the Act is being used as a shield against transparency and is failing to meet its policy objective to foster accountability and trust in our government. Budget 2017 contained no funding for transparency measures and, sadly, there is no direction from the head of the public service regarding transparency, likely meaning there will be minimal impact on the culture of secrecy within the public service. To top it off, institutional performance in relation to compliance with the Act is showing signs of decline, while Canadians' demand for information increases. Comprehensive reform of the Act is essential and long overdue, especially in the face of the expanding information realities of the 21st century. A lot of work needs to be done before this government can meet its transparency promises. (Canada, Office of the Information Commissioner, 2017)

5 For evidence on legislative interest in performance reports, and the link between performance reporting and performance itself at local level, see McDavid (2017).

6 It is important to note that the new Results and Delivery Secretariat also includes two quite distinct components, the results and delivery group headed by an assistant secretary to cabinet, oriented toward appraisal and tracking, and an innovation hub headed by an assistant deputy minister, oriented toward experimentation and adaptive management. (This latter might be seen as distinct from the results group in the same way that the organization division of the Planning Branch was distinct from its evaluation divisions.)

7 See the Treasury Board Directive on Results (Canada, Treasury Board Secretariat, 2016a) for a detailed description of the policy.

8 Two features might be thought striking: the first—central to our message here—is the inclusion of the Minister of Democratic Institutions in expenditure management concerns; the second is the prominent reference to the longstanding question of tax expenditures, a challenge that figures centrally in the issue of access by Parliamentary agents, including the Auditor-General, to government information, as noted later in this chapter.

9 Three reports (Schmitz, 1994; Canada, Treasury Board Secretariat, 2005; Aucoin, 2006) tell the story in interestingly different ways.

10 This vision suggests an analogy with the dual process brain described by Kahneman (2011). In the dual process democracy, System I accumulates the observations and builds the heuristics, intuitions and perceptions on which ongoing decisions about action will be based, but System II tries to discipline these decisions, subjecting them to rational analysis based on a more orderly process of building tested knowledge from

accumulated evidence. (It is tempting, but obviously far-fetched, to think of the role of the Senate in this context as the chamber of 'sober second thought'.)

11 This is also where—as a glance at the comment streams on tracking sites will attest—disputes arise about whether delivery of what was promised has actually happened, whether commitments have truly been fulfilled.

12 As Barry Carin (2015) noted in a 2011 symposium, "We feel like we make decisions, but many of our decisions are made by default, by people who design forms. An excellent example is the choice of default option for the organ donation decision on driver license forms." Carin's insightful but sadly overlooked paper appears in the symposium proceedings edited by Parson (2015). His introduction of this "new paternalism" now reflected in Nudge units stimulated at the time much discussion of the extent to which the values and choices embraced by public servants seep into public policy independently of elected representative governments.

13 It has been argued that a mid-term attempt by the present government to report on promises delivered is not persuasive (Rathgeber, 2017).

14 See the spring 2017 message of the auditor general to Parliament (Canada, Office of the Auditor General, 2017). The May 12, 2017, order-in-council responding to that message does not resolve the deeper ongoing problem; only a full-scale renewal of access to information policy will do that, and such reform should be, as promised, the subject of a much more sincere and forceful effort than is currently evident. (Unfortunately, the April 2016 directive on the subject [Canada, Treasury Board Secretariat, 2016c] does not do the job, because subsection 2.4 leaves wide open the definition of "Cabinet confidences" that lies at the heart of the issue.)

15 On these more ambitious ideas, going beyond open government to open source governance, collective intelligence, and "upstreaming citizen participation," see Lenihan (2012), Landemore (2013) and online portals (*Open Governance Research Exchange*, n.d.) (Hudson, (2016).

16 See Bostrom (2014) and Grace, Salvatier, Dafoe, Zhang, and Evans (2017) for examples. Perhaps a starting point for such machine learning might be found in the Evaluation Checklists at Evaluation Center (2018).

REFERENCES

Aucoin, P. (2005). *Decision-making in government: The role of program evaluation*. (Discussion Paper). Retrieved from the Treasury Board of Canada Secretariat website: https://www.tbs-sct.gc.ca/cee/tools-outils/aucoin-eng.asp

Aucoin, P. (2006). Improving government accountability. *Canadian Parliamentary Review*, *29*(3), 20–26. Retrieved from http://www.revparl.ca/29/3/29n3_06e_Aucoin.pdf

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford, UK: Oxford University Press.

Bourgeois, I., & Whynot, J. (2018). Strategic evaluation utilization in the Canadian federal government. *Canadian Journal of Program Evaluation*, *32*(3), 327–346.

Bourgon, J. (2009). *Program review: The government of Canada's experience eliminating the deficit, 1994–1999: A Canadian case study*. Ottawa, ON: Institute for Government.

Budhwani, S., & McDavid, J. C. (2017). Contribution analysis: Theoretical and practical challenges and prospects for evaluators. *Canadian Journal of Program Evaluation*, *32*(1), 1–24. https://doi.org/10.3138/cjpe.31121

Canada, Office of the Auditor General. (2017, spring). *Message from the Auditor General.* Retrieved from Office of the Auditor General website: http://www.oag-bvg.gc.ca/internet/English/parl_oag_201705_00_e_42222.html

Canada, Office of the Information Commissioner. (2017). Access to Information Act reform: A broken promise? In *Annual report 2016–2017* (Appendix A). Retrieved from Office of the Information Commissioner website: http://www.oic-ci.gc.ca/eng/rapport-annuel-annual-report_2016-2017_11.aspx

Canada, Office of the Prime Minister. (2015, November 12) President of the Treasury Board of Canada Mandate Letter. Retrieved from Office of the Prime Minister website: https://pm.gc.ca/eng/president-treasury-board-canada-mandate-letter

Canada, Privy Council Office. (2016, June 13). *Blueprint 2020 and public service renewal.* Retrieved from Government of Canada website: https://www.canada.ca/en/privy-council/topics/blueprint-2020-public-service-renewal.html

Canada, Task Force on Program Review. (1986). *An introduction to the process of program review*, Ottawa, ON: Minister of Supply and Services.

Canada, Treasury Board. (1968). *Financial management in departments and agencies of the government of Canada*. Ottawa, ON: Treasury Board.

Canada, Treasury Board. (1978). *A-base expenditure review: A report on recent experience in the federal government.* Ottawa, ON: Supply and Services Canada.

Canada, Treasury Board Secretariat. (2005). *Review of the responsibilities and accountabilities of ministers and senior officials.* Retrieved from Treasury Board of Canada Secretariat website: https://www.tbs-sct.gc.ca/report/rev-exa/ar-er-eng.pdf

Canada, Treasury Board Secretariat. (2009). *Evaluation policy suite.* Retrieved from Treasury Board of Canada Secretariat website: https://www.tbs-sct.gc.ca/cee/pol-eng.asp

Canada, Treasury Board Secretariat. (2013–2014). *Evaluation of the 2009 Policy on Evaluation.* Retrieved from Government of Canada website: https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/centre-excellence-evaluation/evaluation-2009-policy-evaluation.html

Canada, Treasury Board Secretariat. (2016a). *Directive on results*. Retrieved from Treasury Board of Canada Secretariat website: https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=31306

Canada, Treasury Board Secretariat. (2016b). *Experimentation direction for deputy heads*. Retrieved from Innovation Hub website: https://www.canada.ca/en/innovation-hub/services/reports-resources/experimentation-direction-deputy-heads.html

Canada, Treasury Board Secretariat. (2016c). *Interim directive on the administration of the Access to Information Act*. Retrieved from Treasury Board of Canada Secretariat website: https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=18310&section=html

Canada, Treasury Board Secretariat. (2016d). *Policy on results*. Retrieved from Treasury Board of Canada Secretariat website: https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=31300

Carin, B. (2015). What is an honest policy analyst to do? In E. A. Parson (Ed.), *A Subtle Balance: Expertise, evidence, and democracy in public policy and governance, 1970–2010* (pp. 222–245). Montreal, QC: Mcgill-Queens University Press.

Chen, K., Bendle, N., & Soman, D. (2017). *Policy by design: The dawn of behaviourally-informed government.* Retrieved from the University of Toronto, Rotman School of Management website: http://www.rotman.utoronto.ca/FacultyAndResearch/ResearchCentres/BEAR/Our-Research/White-Papers-and-Reports

Chouinard, J. A. (2013). The case for participatory evaluation in an era of accountability. *American Journal of Evaluation*, *34*(2), 237–253. https://doi.org/10.1177/1098214013478142

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.

Curran, R. (2016, April 27). Will deliverology work for the federal government? *Policy Options*. Retrieved from Policy Options website: http://policyoptions.irpp.org/magazines/april-2016/is-deliverology-right-for-canada/

Curry, B. (2011, September 20) Flaherty defends $90,000-a-day advice on cutting public service. *Globe and Mail.* Retrieved from https://www.theglobeandmail.com/news/politics/ottawa-notebook/flaherty-defends-90000-a-day-advice-on-cutting-public-service/article617605/

Dobell, A. R. (1999). Evaluation and entitlements: Hartle's search for rationality in government. In R. Bird, M. Trebilcock, & T. Wilson (Eds.), *Rationality in public policy: Retrospect and prospect, a tribute to Douglas G. Hartle* (pp. 79–108). Toronto, ON: Canadian Tax Foundation.

Dobell, R. (2003). *The role of government and government's role in evaluating government: Insider information and outsider beliefs* (Research Paper No. 47). Panel on the Role of Government in Ontario. Retrieved from https://web.archive.org/web/20041103045126/http://www.law-lib.utoronto.ca:80/investing/reports/rp47.pdf

Dobell, R., & Zussman, D. (1981). An evaluation system for government: If politics is theatre, then evaluation is (mostly) art. *Canadian Public Administration*, *24*(3), 404–427. https://doi.org/10.1111/j.1754-7121.1981.tb00341.x

Eliadis, P., Furubo, J.-E., & Jacob, S. (2011). *Evaluation: Seeking truth or power?* New Brunswick, NJ: Transaction.

Evaluation Center. (2018). *Evaluation checklists*. Retrieved from Western Michigan University website: https://wmich.edu/evaluation/checklists

French, Richard D. (2018a). Evidence-based policy: Four schools of thought. Retrieved from http://sciencessociales.uottawa.ca/affaires-publiques-internationales/sites/sciencessociales.uottawa.ca.affaires-publiques-internationales/files/ebp-4_final_.pdf

French, Richard D. (2018b). Lessons from the evidence on evidence-based policy. Retrieved from http://sciencessociales.uottawa.ca/affaires-publiques-internationales/sites/sciencessociales.uottawa.ca.affaires-publiques-internationales/files/lessons_v.2.pdf

Gold, J. (2014). *International delivery: Centres of government and the drive for better policy implementation*. Toronto, ON: Mowat Centre; London, UK: Institute for Government. Retrieved from https://mowatcentre.ca/wp-content/uploads/publications/97_international_delivery.pdf

Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2017). *When will AI exceed human performance? Evidence from AI experts*. Retrieved from Cornell University Library website: https://arxiv.org/pdf/1705.08807.pdf

Graeff, E. (2017). *Designing for monitorial citizenship*. Retrieved from The Impacts of Civic Technology Conference website: http://tictec.mysociety.org/2017/presentation/designing-monitorial-citizenship

Hartle, D. G. (1973). A proposed system of program and policy evaluation. *Canadian Public Administration*, *16*(2), 243–266. https://doi.org/10.1111/j.1754-7121.1973.tb01609.x

Hodge, R. A. (2017). *Towards contribution analysis* (WIDER Working Paper 2017/19). Retrieved from Wider University website: https://www.wider.unu.edu/sites/default/files/wp2017-19.pdf

Hudson, A. (2016, January 21). The value of open governance [Web log post]. Retrieved from Global Integrity website: https://www.globalintegrity.org/2016/01/the-value-of-open-governance-adaptive-learning-and-development/#comment-1029

Kahneman, D. (2011). *Thinking fast and slow*. Toronto, ON: Anchor Canada.

Landemore, H. (2013). *Democratic reason: Politics, collective intelligence, and the rule of the many*. Princeton, NJ: Princeton University Press.

Landemore, H. (2016). Beyond the fact of disagreement? The epistemic turn in deliberative democracy. *Social Epistemology*, *31*(3), 277–295. https://doi.org/10.1080/02691728.2017.1317868

Lenihan, D. (2012). *Rescuing policy: The case for public engagement*. Toronto, ON: Public Policy Forum.

Longo, J. (in press). Digital tools for rapid policy design. In M. Howlett & I. Mukherjee (Eds.), *Routledge handbook of policy design*.

Maxwell, N. (1986). Linking ongoing performance measurement and program evaluation in the Canadian federal government. *Canadian Journal of Program Evaluation*, *1*(1), 39–46. Retrieved from https://evaluationcanada.ca/secure/01-1-039.pdf

Mayne, J. (1986). In defense of program evaluation. *Canadian Journal of Program Evaluation*, *1*(2), 97–102.

Mayne, J. (2018). Linking evaluation to expenditure reviews: Neither realistic nor a good idea. *Canadian Journal of Program Evaluation*, *32*(3), 316–326.

McDavid, J. (2017). Local government performance management: What we know, what we suspect and what we still need to find out. Retrieved from Canadian Association of Government Finance Officers website: http://cagfo.ca/wp-content/uploads/Local-Government-Performance-Management-Dr.-Jim-McDavid-Sep-14-2017.pdf

McDavid, J. C., Huse, I., & Hawthorn, L. R. L. (2012). *Program evaluation and performance measurement: An introduction to practice* (2nd ed.). Thousand Oaks, CA: Sage.

Open Governance Research Exchange. (n.d.). Retrieved from http://govlab.github.io/open-government-research-exchange

Parson, E. A. (Ed.). (2015). *A subtle balance: Expertise, evidence, and democracy in public policy and governance, 1970–2010*. Montreal, QC: McGill-Queen's University Press.

Pressman, J. L., & Wildavsky, A. (1973). *Implementation: How great expectations in Washington are dashed in Oakland; or, Why it's amazing that federal programs work at all. . . .* Berkeley, CA: University of California Press.

Promise tracker: *Data collection for civic action*. (n.d.). Retrieved from http://promise-tracker.org/

Rathgeber, B. (2017, November 15). Trudeau's mandate tracker: Governing is easy! *iPolitics*. Retrieved from iPolitics website: https://ipolitics.ca/2017/11/15/trudeaus-mandate-tracker-governing-easy/

Savoie, D. J. (2013). *Whatever happened to the music teacher? How government decides and why*. Montreal, QC: McGill-Queens University Press.

Schmitz, G. (1994). *Re-organizing government: The reform debate in perspective*. Ottawa, ON: Library of Parliament. Retrieved from https://lop.parl.ca/Content/LOP/ResearchPublicationsArchive/bp1000/bp376-e.asp

Segsworth, R. V. (2005). Program evaluation in the government of Canada: Plus ça change. . . . *Canadian Journal of Program Evaluation*, *20*(3), 175–197.

Sevinc, K. (2017). *Democracy 4.0*. Retrieved from Kader Sevinc website: http://kader-sevinc.blogactiv.eu/2017/04/10/democracy-4-0-renovation-of-politics-and-smart-citizenship-initiative/

Shepherd, R. P. (2012). In search of a balanced Canadian federal evaluation function: Getting to relevance. *Canadian Journal of Program Evaluation*, *26*(2), 1–45. Retrieved from https://evaluationcanada.ca/system/files/cjpe-entries/26-2-001.pdf

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.

Tasker, J. P. (2017, October 11). Liberal government tells CRA to back off plan to tax employee discounts. *CBC News*. Retrieved from http://www.cbc.ca/news/politics/liberal-government-cra-employee-discount-1.4349651

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth and happiness*. New Haven, CT: Yale University Press.

TrudeauMetre. (n.d.). Retrieved from https://trudeaumetre.polimeter.org/

United States, National Science and Technology Council. (2017, January 20) *About SBST*. Retrieved from SBST website: https://sbst.gov/

Weiss, C. H. (1972). *Evaluation research: Methods of assessing program effectiveness*. Englewood Cliffs, NJ: Prentice-Hall.

Weiss, C. H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, *19*(1), 21–33. https://doi.org/10.1177/109821409801900103

Weiss, C. H. (1999). The interface between evaluation and public policy. *Evaluation*, *5*(4), 468–486. https://doi.org/10.1177/135638909900500408

Zilio, M. (2016, September 30). Liberals spend $200,000 for advice on delivering on campaign pledges. *Globe and Mail*. Retrieved from Globe and Mail website: https://beta.theglobeandmail.com/news/politics/liberals-spend-200000-for-advice-on-delivering-campaign-pledges/article32187629/?ref=http://www.theglobeandmail.com&

Zussman, D. (2017, April). The public sector integrity commissioner (PSIC) in perspective. *Canadian Government Executive*. Retrieved from https://canadiangovernmentexecutive.ca/the-public-sector-integrity-commission-psic-in-perspective/

ABOUT THE AUTHORS

Rod Dobell, BA, MA, (UBC), PhD (MIT), taught economic theory at Harvard before returning to Canada as Professor of Political Economy at the University of Toronto. He has served as Special Advisor (Long Range Economic Planning) to the Deputy Minister of Finance and as Deputy Secretary to the Treasury Board in the Government of Canada, Director of the General Economics Branch at OECD, Director of Research for two Parliamentary Task Forces, and President of the Institute for Research on Public Policy, 1984–1991. In 1991, he was named as the first appointment to the Francis G. Winspear Chair for Research in Public Policy at the University of Victoria, where he is now Senior Research Associate at the Centre for Global Studies.

David Zussman is an adjunct professor at the School of Public Administration at the University of Victoria and Senior Fellow at the Graduate School of Public and International Affairs (GSPIA) at the University of Ottawa. In academia, he held the Jarislowsky Chair in Public Sector Management at the University of Ottawa from 2006 to 2015, was a professor in the School of Public Administration at University of Victoria, and served as Dean of the Telfer Faculty of Management and Director of the GSPIA at the University of Ottawa. Earlier in his career, he was President of the Public Policy Forum, Assistant Secretary to the federal Cabinet for Machinery of Government and Commissioner of the Public Service Commission. He has also had a regular public policy column in the *Ottawa Citizen* and *Canadian Government Executive*.

Appendix 1: Six Questions for Program Review (http://www.gouvernance.ca/publications/96-15.pdf)

1) **Public Interest Test:** Does the program area or activity continue to serve a public interest?
2) **Role of Government Test:** Is there a legitimate and necessary role for government in this program area or activity?
3) **Federalism Test:** Is the current role of the federal government appropriate, or is the program a candidate for realignment with the provinces?
4) **Partnership Test:** What activities or programs should or could be transferred in whole or in part to the private or voluntary sector?
5) **Efficiency Test:** If the program or activity continues, how could its efficiency be improved?
6) **Affordability Test:** Is the resultant package of programs and activities affordable within the fiscal restraint? If not, what programs or activities should be abandoned?