Does “Deliverology” Matter?

by Mark Schacter
June 2016

Does “Deliverology” Matter?

Published by:

Mark Schacter Consulting
Ottawa, Ontario, Canada

www.schacterconsulting.com
mark@schacterconsulting.com

cover photo:
“Tissue Delivery” by Kamyar Adl
used under Creative Commons License 2.0
https://creativecommons.org/licenses/by/2.0/

Table of Contents

Does "Deliverology" Matter?

"Goodhart's law": When a measure becomes a target, it ceases to be a good measure.[1]

A Introduction

In the 1960s the "British invasion" brought the Beatles and other British bands to North America. It changed forever the way we appreciate popular music.

Fifty years on, Canadian public servants are witnesses to another British invasion by "deliverology", an approach to public management first implemented by the British government in the early 2000s.

When the Beatles arrived many music critics were unimpressed. "Not even their mothers would claim they sing well," wrote one reviewer. Another said, "If you don't think about them, they will go away, and in a few more years they will probably be bald…."[2]

Those forgotten critics are reminders of the odds against correctly predicting the success of a brash newcomer bent on upsetting the *status quo*. Nevertheless, I can't help but wonder whether deliverology will turn out to be a major event in Canada, or end up as a flash in the public-management pan. (As another 1960s music critic said, "The Beatles will pass. The question is, what next?")[3]

Deliverology's creator and leading proponent, Michael Barber, has been invited by Canada's newly-elected government to participate in two Cabinet retreats this year. His ideas are influencing Canada's new leaders, and some elements of his system have already been put in place. The full extent to which deliverology will be implemented in Canada remains to be seen. In the meantime, I ask:

Does deliverology matter to the practice of public management in Canada? This leads to four sub-questions:

- Is deliverology really something new?

- Is it relevant to a problem that needs to be fixed?

---

[1] https://en.wikipedia.org/wiki/Goodhart's_law
[2] http://articles.latimes.com/2014/feb/09/opinion/la-oe-beatles-quotes-20140209
[3] *Idem.*

- Does it work?

- Might it do harm?

B Is Deliverology Something New?

The word "deliverology" is new to Canadian public servants. But is the underlying idea new?

Barber describes deliverology as an approach to "managing and monitoring the implementation of activities that have a significant impact on outcomes"[4] or alternatively, "a systematic process for driving progress and delivering results in government and the public sector."[5]

Barber ran the Standards and Effectiveness Unit in Great Britain's Department for Education and Employment between 1997 and 2001 under Prime Minister Tony Blair. From 2001 to 2005 he reported directly to Blair as head of the Prime Minister's Delivery Unit (PMDU). It was in this capacity that Barber systematized the methodology now known as "deliverology".[6] The essential elements of deliverology are:

- establishing a "delivery unit" – i.e. "a small group of dedicated individuals focused exclusively on achieving impact and improving outcomes"[7];

- creating a clear vision of program success;

- quantifying the vision of success through performance measures and targets;

- closely monitoring performance against targets and taking action to address variances; and

- ensuring that there are consequences for good and bad performance.

Putting aside for a moment the dedicated "delivery unit", none of these ideas are new to program management in the Canadian public service. But if the British experience is any

---

[4] The concept is summarized in "Deliverology: From idea to implementation," by Michael Barber, Paul Kihn and Andy Moffit, McKinsey & Company: 2011. http://www.mckinsey.com/industries/public-sector/our-insights/deliverology-from-idea-to-implementation.

[5] *Deliverology 101: A field guide for educational leaders*, by Michael Barber, Andy Moffitt and Paul Kihn, SAGE Publications: 2011, from the Preface.

[6] Following his tenure as an adviser to Blair, Barber became a partner at the consulting firm McKinsey, where he headed its Global Education Practice. In 2011 Barber became Chief Education Adviser at Pearson, a British publishing and education.

[7] "Deliverology: From idea to implementation."

indication, what would be new to Canadian public servants is the intense, single-minded focus on setting targets and monitoring performance against them.

The concept of a dedicated delivery unit would also be new. But the novelty is about form, not substance. That there should be staff focused on creating a vision of what good performance looks like, using data and reporting to track performance and hold people accountable is not new. In Canada this responsibility rests with program managers and their supervisors, and so on up the line to senior executives. Deliverology calls for creating a separate group of people outside the operational chain of command whose job is to challenge program managers and "ask difficult questions" about performance.

From the deliverology perspective it would probably be argued that the dedicated delivery unit is the key difference-maker; that it is only through establishment of such a unit that chronic failures of implementation can be overcome. This is a critical assumption to which I will return.

C Is Deliverology Relevant to a Problem that Needs to be Fixed?

If deliverology is the solution, then what is the problem?

A summary of deliverology[8] states that governments and public agencies "set ambitious reform goals" and plans to achieve them, but that such plans "frequently … fall by the wayside and reform goals remain unmet." Public-sector organizations often have difficulty finding "ways to define and execute their highest-priority objectives so that they have the greatest possible impact."

In other words, the general problem addressed by deliverology is that governments are prone to making big announcements and failing to deliver fully on them.

Targets

The deliverology apparatus – the delivery units, the establishment of performance "trajectories", the "monthly notes" and quarterly meetings, etc. – is built around the creation and validation of quantified targets, and then monitoring performance and making program-management decisions in relation to those targets. Barber describes performance targets as one of deliverology's "most effective tools".

This has obvious appeal. It sounds clear, simple, data-driven and scientifically rigorous.
But a strong argument can be made that targets are an unreliable – and sometimes entirely unsuitable – instrument for strengthening public program delivery. A well-documented body of

[8] "Deliverology: From idea to implementation."

experience leads one to question the relevance of targeting as a primary instrument for driving performance improvements in the public sector. To the contrary, there is plenty of evidence to suggest that intensive use of targets is as likely to compound as to alleviate problems of public sector performance – an issue I will revisit further on.

Much of the cautionary evidence about targeting comes from Barber's own country and is drawn from the period when he was running the PMDU. A report in 2003 by a British parliamentary committee[9] expressed skepticism about the capacity of targets to improve public sector performance. "We discovered that targets and results were different things," the report concluded.

The world of promises, announcements and desired results is inhabited by elected officials, while the world of delivery belongs to public servants. Interaction between the two worlds is fraught with incompatible incentives and motivations; full delivery of commitments made by elected officials is a multifaceted and highly nuanced problem.

Deliverology responds with a simple, mechanistic solution. Its implicit assumption is one I expect public servants would find as offensive as it is incorrect, i.e. the principal causes of failure to deliver on politicians' goals *can always be traced to shortcomings in the public service*. How else can one explain the heavy emphasis on targets as both carrots and sticks that will push public servants toward faster, fuller complete delivery?

Barber's choice of words is telling. He says a key function of a delivery unit is "taking any excuses off the table … [because] the tendency of any system is toward inertia"; regular meetings initiated by delivery units to review performance against targets will "create a sense of urgency"; delivery unit staff will be "top talent" – "drawn from among the most talented and qualified people inside or outside the system."[10]

Speaking at a World Bank panel discussion on public management, Tony Blair said he was motivated to implement deliverology when he discovered that "the traditional British civil service has never been about delivering results."[11]

The implication: average public servants are mediocre performers (in comparison to "top talent" skimmed off the top for delivery units); if left to their own devices, they will make excuses, lack a sense of urgency, and, above all, fail to "deliver".

---

[9] "On Target? Government By Measurement," House of Commons Public Administration Select Committee, Fifth Report of Session 2002-03, Volume I.
[10] "Deliverology: From idea to implementation."
[11] http://www.worldbank.org/en/news/speech/2013/04/10/delivering-results-conversation-jim-yong-kim-tony-blair-michael-barber

It is of course true that causes of disappointing delivery will sometimes be found in the public service. But it is simplistic to make this the entire story. Delivery shortfalls are just as likely to be rooted in the political level. Politicians are prone to making promises without a clear understanding of what it would take to deliver on them, or of a Department's capacity to deliver, or of realities "on the ground" that are only understood by program staff. On top of all this, there is a range of impediments to performance that are beyond the control of *either* politicians *or* public servants.

Deliverology is technical solution that ignores the complex reality and inherent ambiguity of public service and public management. It assumes a narrow and specific delivery problem that can be traced to the bureaucracy. It assumes that performance targets are appropriate instruments for addressing all manner of delivery problems. If these assumptions are wrong, then deliverology is an elaborate solution in search of a problem.

D Does Deliverology Work?

Barber himself of course claims that deliverology is effective; so does his former consulting company, McKinsey. Tony Blair also praises the value of deliverology. These views cannot be considered objective[12].

Is there independent evidence that deliverology works? It is noteworthy that the President of the World Bank, Jim Yong Kim, has sought the advice of Barber and Blair with regard to strengthening public service delivery in developing countries. And dozens of leaders around the globe have signed on to deliverology. A recently published biographical summary of Barber says he has been invited to work on "education reform and public service reform in 49 countries."[13]

Such expressions of confidence are persuasive. But *belief* that deliverology works is not equivalent to evidence that it is effective. Even though deliverology is founded on rigorous and relentless devotion to objective empirical evidence, *it is difficult to find persuasive, independent evidence that it actually contributes in a significant way to improved public service delivery*.

A study published in 2006[14] examined the impact of deliverology-inspired performance targets on healthcare service delivery in the England between 2001 and 2005 (coincident with the

[12] Tony Blair's use of his connection with deliverology as a basis for promoting his consulting business has been widely reported. For example, see http://www.telegraph.co.uk/news/politics/tony-blair/11547808/Revealed-how-Tony-Blair-makes-his-millions.html; http://www.dailymail.co.uk/news/article-3045658/Tony-Blair-s-bizarre-deal-sell-deliverology-Colombia-paid-United-Arab-Emirates.html and http://foreignpolicy.com/2013/10/23/tony-blairs-excellent-balkans-adventure/

[13] "The Good News from Pakistan," by Michael Barber, *REFORM*, March 2013, p. 2 http://www.reform.uk/publication/the-good-news-from-pakistan/

[14] "Have Targets Improved Performance in the English NHS," by Gwyn Bevan and Christopher Hood, *British Medical Journal*, Vol. 332, No. 7538 (Feb. 18, 2006), pp. 419-422.

period when Barber was head of the PMDU). It found that "reported performance"[15] in relation to key targets had improved, but that the targeting regime also caused significant negative effects felt by both providers and recipients of healthcare. The researchers were unable to conclude that the net impact of deliverology was positive. "… (Q)uestions remain over the extent to which improvements in targeted performance in the English NHS [National Health Service] were undermined by … gaming …"[16]

A study published in 2012 examined the impact of wait-time targets on Emergency Departments (EDs) in English hospitals during the period 2003-06.[17] The findings were inconclusive. On the one hand, targets led to shorter wait times. Furthermore,

> The existence of a time target … focused attention on EDs and their patients. Many departments reported receiving more resources to meet the targets. Traditional ways of working have been challenged and innovative strategies implemented.

But the fact that targets were met did not lead to any improvement in the overall patient experience.

> We hoped that the target would have led to improved processes, resulting in patients being treated sooner and leaving earlier … without diminishing time for physician-patient interactions and care. We did not observe this pattern …. *EDs are performing to the target but not improving overall care*[18].

A study published in 2006 reached a similar conclusion regarding police services in England and Wales.[19] The article described the overwhelming influence of "pressure now exerted by the Prime Minister's Public Service Delivery Unit". It observed that

> … 'performance' was seen to relate to targets and the collection of quantitative data on crime statistics, complaints, sickness rates, charge and caution rates. Yet the data required did not refer to quality of service …

---

[15] The reference to *reported* performance is significant. In any rigorously enforced targeting regime where staff feel strong pressure to meet targets, it is inevitable that some reporting will be fabricated to "demonstrate" that targets have been met. As the authors observed, the targeting system used in the English health system was "criticised for … similarities to the target regime of the former Soviet Union, although NHS managers were threatened with loss of their jobs rather than their life or liberty." *Ibid.,* p. 420.

[16] *Ibid.*, p. 421.

[17] "Time Patients Spend in the Emergency Department: England's 4-Hour Rule – A Case of Hitting the Target but Missing the Point?", by Suzanne Mason et. al., *Annals of Emergency Medicine*, Vol. 59, No. 5, May 2012, pp. 341-349.

[18] Emphasis added.

[19] "Policing performance: The impact of performance measures and targets on police forces in England and Wales," by Barry Loveday, *International Journal of Police Science and Management,* Vol. 8, No. 4, pp. 282-293

Claims that deliverology caused significant improvements in the education sector of Pakistan's Punjab province have also raised questions. Barber began work on education reform in Punjab in 2011. Two years later he reported that his team's efforts had produced:

- nearly 1.5 million additional children enrolled in school;
- an increase in facilities with functioning electricity, drinking water, toilet and boundary walls from 69 to 91 percent;
- an increase in student attendance from 83 percent to 92 percent;
- an increase in teacher attendance from 81 pe cent to 91 percent;
- 81,000 new teachers hired on merit;
- simple, easy-to-use lesson plans for every teacher and new textbooks for every student;
- evidence of improving learning outcomes;
- implementation of a voucher scheme enabling over 140,000 out-of-school children of poor families to attend private schools.[20]

This interpretation of events was disputed by Jishnu Das, a research economist at the World Bank and Visiting Fellow at the Centre for Policy Research in New Delhi with a decade of experience researching educational outcomes in Punjab. He observed that:

- data on school enrolment in Punjab for the period to which Barber refers were ambiguous; it was difficult to determine what was really happening to enrolment while Barber was active there;
- there were impressive improvements in school inputs – facilities and teacher attendance – during the period of Barber's involvement but these could have been a continuation of a strong positive trend that began as early as 2003 (eight years before Barber arrived); the "trend under deliverology looks much like the trend before it", Das said; and

[20] "The Good News from Pakistan."

- student performance in Punjab showed "impressive gains", but gains in the neighboring province of Khyber Pakhtunwa – where Barber was not involved – were even larger.

Das' conclusion: "The lack of credible public data and the lack of third-party evaluations of the program make it difficult to go to bat for the deliverologists at this time." [21]

E Might Deliverology Do Harm?

The biggest "open secret" in the public sector is the awkward phenomenon of fiscal-year-end spending. Everyone who manages a budget faces the same performance target: a budget balance as close as possible to zero by the end of the year. Managers with uncomfortably large balances in the last fiscal quarter find reasons to spend money that might not have motivated expenditure a few months earlier.

There's nothing illegal about this – perhaps nothing unethical either – but it is all rather uncomfortable. Why? Because the main factor driving decision-making is an administrative target rather than consideration of the public interest. Public servants respond rationally to the message conveyed by the target – "unspent money at the end of the year indicates incompetent financial management".

Now imagine what it would be like to have a significant proportion of all program decision-making driven by targets throughout the year. This was deliverology at its peak under Michael Barber and Tony Blair in Great Britain in the early 2000s

Theory strongly suggests and experience has shown that targets are rarely a sound basis for decision-making and performance measurement in the public sector. Public management scholarship has identified the mechanisms through which targets generate counter-productive behavior[22]. They include:

- Synecdoche. A part is assumed to stand for the whole, as when meeting a target such as waiting no more than four hours in a hospital emergency room is taken to represent the totality of a patient's experience. This will cause staff and managers to focus on the target at the expense of other, perhaps more important, issues. Hence, "meeting the target and missing the point".

- Threshold effects. A minimum-standard target may motivate weak performers to improve, but may also de-motivate superior performers who were already above the

[21] "A Data Guide to Sir Michael Barber's 'The Good News from Pakistan'", by Jishnu Das. http://blogs.worldbank.org/developmenttalk/data-guide-sir-michael-barber-s-good-news-pakistan
[22] "The logics of performance management," by Christopher Pollitt, *Evaluation*, Vol. 19, No. 4, 2013, pp. 346-363.

threshold.

- Ratchet effects. If next year's targets are based on last year's performance then managers may intentionally under-perform so as not to set themselves too difficult a task for the following year.

- Definitional drift. Measurement criteria may be "stretched" (redefined) to make it easier to meet a target, as when hospitals redefine a stretcher as a "mobile bed" to meet a target for admitting patients within a time limit.

- Cheating. When judgments about performance – and related rewards/punishments – are tied tightly to targets, it is reasonable to expect deliberate mis-reporting and other types of rule-breaking.

The public-management literature – most notably, accounts of the British public service during the early 2000s – is rife with descriptions of these phenomena. Here are a few examples:

- To meet the time target for getting a first responder to the scene of a life-threatening emergency, the Ambulance Service introduced lay first responders (as opposed to fully qualified paramedics). It was questionable whether the lay responders were able to handle the full range of emergencies to which they responded.[23]

- Wait-time targets for new outpatient appointments at an eye hospital were achieved by cancelling or delaying follow-up appointments. As a result, at least 25 people lost vision over a two-year period.[24]

- Physicians' offices responded to a target that 100 percent of patients be offered an appointment within two working days by refusing to book any appointments more than two days in advance.[25]

- A school teacher helped pupils cheat on standardized tests to avoid negative consequences for the school of low scores.[26]

[23] "On Target? Government by Measurement," p. 18
[24] *Idem*.
[25] "Have targets improved performance in the English NHS," p. 420.
[26] *Ibid.,* p. 19

- Police commanders in low-crime areas were judged to be performing poorly when they failed to reduce what were already low crime rates.[27]

- To boost a police force's performance related to targets for responding to emergency calls, members of the force made bogus emergency calls.[28]

I assume proponents of deliverology would say that these examples do not discredit performance targeting; they simply underscore the importance of creating good targets and managing the performance measurement process properly. *But this response fails to account for the fundamental incompatibility of targeting with much of what the public sector actually does.*

The assumption behind a quantitative performance target is that a description of good performance can be expressed in a single number. For this to be true you need a production process that is highly standardized, fully specified and more or less invariable. In other words, you need to be doing the same thing over and over again in exactly (or almost exactly) the same way. Under these circumstances, managers know *exactly* what to expect and can therefore set precise performance targets. Many industrial processes are like this – for example, a steel mill or an automobile assembly line. We know exactly how many cars should roll off the line, or how much steel should be produced, in a given period of time. Failure to meet a target is a signal that something unusual is happening, and prompts the manager to investigate further. *In other words, the target conveys meaningful information.*

Public programs are not delivered through industrial processes, although some come close. Processing tax returns and issuing income support payments are examples of quasi-industrial processes to which performance targeting may be suitable. If we have a good idea of how many tax returns or income support payments a clerk should be able to process in a work day, then it may be reasonable to set performance targets.

But much of what the public sector does is not like processing tax returns; methods may vary from situation to situation; judgment may have to be exercised; multiple stakeholders may have to be satisfied and multiple objectives served simultaneously. Complexity and ambiguity rule. And as Britain's experience has shown, even service delivery processes that initially appear repetitive and standardized – such as managing patients in an emergency department – turn out to be stubbornly complex.

Targets do not convey meaningful information under complexity and ambiguity. Worse, they send misleading signals. Consider for example a program run by the Department of Canadian

[27] "Policing performance: The impact of performance measures and targets on police forces in England and Wales," p. 286.
[28] *Ibid.,* p. 289.

Heritage that provides "grants and contributions for infrastructure projects for professional, not-for-profit arts and heritage organizations and municipal and provincial government and agencies with a mandate for arts and heritage, and equivalent Aboriginal peoples' organizations."[29] Suppose an efficiency target were set for the program along the lines of "cost per $ million in grants/contributions issued".[30]

Proposed projects might be small or large, simple or complex, of short or long duration. Proponents might range from sophisticated provincial government organizations experienced in submitting funding proposals to small NGOs and Aboriginal organizations that require significant assistance to complete their applications. Suppose it took 30 person-days of Departmental time to process and make the first disbursement on a proposal for a $100,000 grant from a two-person Aboriginal organization in a remote community, and 25 person-days to do the same on a proposal from a provincial government agency for a $500,000 grant. Management of the second proposal appears to be *six times more efficient* than the first (50 person-days/$ million vs. 300 person-days/$ million).

But isn't it the case that working with the remote Aboriginal organization *should be* more time-intensive than working with the provincial organization? If efficiency is expenditure per unit of output, then the outputs in this example are not comparable. Issuing a grant to the provincial organization is fundamentally different from issuing a grant to the Aboriginal organization. And what if – as is sometimes the case – a separate objective of the program was to strengthen planning and management capacity of community arts organizations via time-intensive support from program staff for completing applications? There is no "normal" here, no plausible efficiency target.

Another example: a program run by Environment and Climate Change Canada that aims to improve the situation of species at risk. Suppose one were to have a target for reducing the number of species on the "at risk" list by a given number every year. Two factors undermine the value of the target: i) many variables outside the control of the program affect the health of species; and ii) the target treats every species as being of equal importance from an ecological perspective. Suppose that the program missed its target, but that one species considered among the most important in Canada in terms of preserving biodiversity was taken off the "at risk" list? Wouldn't that be "success"?

These examples highlight how targets – the primary instruments used by deliverology – may bias our understanding of performance and lead to misguided decisions in cases of particular

[29] This description of the Canada Cultural Spaces Fund is taken from Canadian Heritage's 2015-16 Report on Plans and Priorities.

[30] Guidance from the Treasury Board Secretariat suggests this type of efficiency measure for grants and contributions programs.

programs. I want to extend the argument now to two important system-wide issues: i) managing for results; and ii) fostering innovation.

Results

As described above (p. 2) Barber equates "delivery" with achievement of results or outcomes. In *Deliverology 101* he offers a "systematic process for … delivering results".

In Barber's work I cannot find a precise definition for "outcomes" or "results". I will follow accepted usage in the Government of Canada where:

- "outcome" and "result" mean the same thing and are used interchangeably; and
- the words refer to desired social, economic or environment changes that are influenced by the delivery of programs.

In other words, programs *deliver* what in the Canadian system are referred to as "outputs", i.e. money, information, laws & regulations, services. Programs *control* delivery of outputs. Outputs *influence* the occurrence of outcomes. For example:

- a program delivers a grant (output) to a community-based cultural organization;
- cultural organization uses the grant to fund a local theatre company (outcome);
- the theatre company produces theatrical productions (outcome);
- the cultural life of the local community is enriched (outcome).

It's important to recognize that the only thing *delivered* by the government in this scenario is the grant (the "output"). Everything else is a matter of influence. Programs – and governments as a whole, for that matter – *do not deliver results*. They only deliver outputs, in the expectation that outputs will influence outcomes. Barber's language – the suggestion that his system can help governments do better at delivering *results* – implies a fundamental misunderstanding of what governments do, and indeed are capable of doing.

But perhaps Barber meant to refer to delivery of what we call "outputs". This would be equally unpalatable. Since the late 1990s the Treasury Board Secretariat has focused – with good reason – on encouraging Departments to manage for results rather than outputs. The underlying idea is that output production, on its own, is worthless. Of what value is the grant to the cultural organization if it doesn't contribute to strengthening the organization's capacity

so that it can fund the theatre company so that the cultural life of the community can be enriched?

So if deliverology – through its target-setting and tracking mechanisms – is about getting managers to focus single-mindedly on production of outputs, then its adoption would represent an unfortunate reversal of direction for public management in Canada. It would signal a return to favoring production of outputs over contribution to outcomes.

Innovation

Canada's newly-elected federal government has identified fostering innovation as a high priority. It aims for a "confident nation of innovators".[31] Making this a reality will require innovative thinking and behavior within as well as outside the public service – we will need a "confident public service of innovators".

Can innovation thrive in an organizational culture characterized by adherence to targets? The British experience suggests strongly that it cannot. To the extent that deliverology generated creative energy among public servants, it was expended on finding ways to beat – through gaming and outright cheating – a system perceived as detached from the realities of public management.

> The Government hopes that target setting will encourage service providers to apply creativity in making their activities contribute effectively to delivery. But in some cases creativity is being directed more to ensuring that the figures are right than to improving services. This is where measurement ceases to be a means to an end and becomes and end in itself.[32]

That a target culture is likely to be incompatible with innovation should not come as a surprise. Management by target is about conformity, predictability and meeting a minimum standard. Innovation is about diversity, uncertainty and breakthroughs – while accepting that many efforts will end in failure.

> [Deliverology will] encourage conformity rather than risk-taking. This is primarily because the risk of failure and what follows on from that is seen as outweighing any potential reward for success.[33]

---

[31] "Building an inclusive and innovative Canada," news release, June 14, 2016 http://news.gc.ca/web/article-en.do?nid=1084439

[32] "On Target? Government by Measurement," p. 17

[33] "Performance Management and the Decline of Leadership within Public Services in the United Kingdom," by Barry Loveday, *Policing*, Vol. 2, No. 1, pp. 120-130.

F Conclusion

I have broken down the question "does deliverology matter?" into four sub-questions: i) is deliverology new? ii) is it relevant to a problem? iii) does it work? and iv) might it do harm?

Is it new? It is a new way of packaging ideas that should already be familiar to Canadian public servants. Its most striking elements are a sharp focus on setting targets and monitoring performance against them, and the introduction of "delivery units" outside the operational chain of command. Delivery units have an important role to play in creating targets, monitoring performance information, provoking management responses to performance information, providing expert support to operational managers and generally "asking difficult questions".

Is it relevant to a problem? It is difficult to see how deliverology is relevant to a problem in the Canadian public service unless you believe that Canadian public servants are generally unable and/or unwilling to focus on "delivery", and will only do so when subjected to a system with the features of deliverology. And even if there were a delivery problem in the public service, a large body of theory and empirical evidence suggests that an emphasis on targets is not a relevant solution.

Does it work? There is no persuasive, independent evidence I was able to find that suggests deliverology makes a significant positive difference to the quality of public management.

Might it do harm? Yes, in at least three ways:

- Imposition of a performance management system with heavy emphasis on targets risks distorting behavior of program managers and executives toward meeting targets while ignoring contributions to outcomes that affect the lives of Canadians. Targets are only relevant management tools for some aspects of what the Canadian public sector does (e.g. quasi-industrial processes like handling income tax returns). Universal application of targets across the public service would be inappropriate and harmful.

- Emphasis on "delivery" risks focusing the attention of program managers and executives on outputs (what programs deliver) instead of outcomes (social, economic or environmental benefits experienced by Canadians).

- A heavy emphasis on performance targets risks stifling innovation and creativity in the public service.

Deliverology, like any public-management system, is merely an instrument; a means to an end. It is not a magic formula, nor does it have inherent value. Its value, if any, is to be found in its capacity to strengthen the quality of public management.

"Strengthening the quality of public management" means, in my view, building the capacity of the public service to design and deliver programs (i.e. deliver outputs) that contribute to social, economic or environmental outcomes that improve the lives of Canadians. There is no independent evidence that deliverology is an effective way to achieve this.

Certain aspects of deliverology are valuable and relevant to the Canadian context – notably, the emphasis on defining outcomes and being clear about how success will be measured. These principles are already embedded in the Canadian public sector's management policies.

So deliverology *does* matter to public management in Canada, but not in the way that its proponents would suggest. It matters not because it is relevant to a Canadian problem and not because it is a proven public-management approach, but rather because it may cause significant harm to Canada's public service should the government choose to implement all the main features of Barber's system.

About the Author

Mark Schacter is a public-management consultant and author based in Ottawa, Canada. For more than 20 years he has advised governments, NGOs and international organizations on practical and policy issues in governance, accountability, performance measurement, evaluation, risk and strategy. He is also an accomplished trainer, having designed and led workshops on topics such as governance, performance measurement, evaluation, strategic planning and risk management.

In the 1990s Mark was on the staff of the World Bank at its headquarters in Washington DC; he worked in Africa, Asia and Eastern Europe to help strengthen the governance and accountability of major public institutions, and to support reforms in private sector development. Subsequently, he occupied senior positions at the Institute On Governance and the Conference Board of Canada, both in Ottawa, before launching his consulting firm, Mark Schacter Consulting, in 2003.

Mark has a BA from Yale University, an LL.B. from Oxford University and an MBA from the University of Ottawa.