Burt Perrin

Bringing accountability up to date with the realities of public sector management in the 21st century

Abstract: This article identifies basic shortcomings of traditional approaches to accountability and considers some of the reasons for the persistence of an approach that is known to: provide an inaccurate and distorted view of actual performance; inhibits rather than facilitates improved performance; and contributes to less rather than more confidence in government. The article then presents a vision of accountability more in keeping with the realities of public sector management in the twenty-first century.

Sommaire : Le présent article indique les principales déficiences dans les façons traditionnelles d'aborder l'obligation de rendre compte et examine certaines des raisons pour lesquelles on continue de recourir à cette démarche qui, on le sait, entraîne les conséquences suivantes : cela fournit une fausse et inexacte représentation de la performance réelle, empêche plutôt que ne favorise une meilleure performance, et contribue à avoir moins plutôt que davantage confiance dans le gouvernement. L'article présente ensuite une vision de l'imputabilité qui correspond mieux aux réalités de la gestion du secteur public au XXI[e] siècle.

Introduction

In recent years, there has been a major shift in public sector management from a focus on activities and processes to a focus on *results* – benefits to citizens arising from government processes. There also is increasing recognition that government interventions, by their very nature, are complex and operate in uncertain environments where adaptability is needed rather than sticking to a plan that can quickly become out of date. As well, significant outcomes rarely follow from just a single action or actor.

These shifts have implications for *all* aspects of public sector management, including accountability. But, as this article documents, traditional approaches to accountability, designed primarily for checking compliance with rules and procedures and expenditures against budgets, have not

Burt Perrin is an independent consultant based in France who assists governments and other organizations internationally with advice about planning, evaluation management, design and quality assurance (Burt@BurtPerrin.com). He would like to thank Harry Jones, John Mayne, Nicoletta Stame, Jim McDavid, and Alison Perrin for many helpful suggestions to an earlier draft of this article. In particular, he would like to thank Jim McDavid for his encouragement, support, and editorial assistance..

CANADIAN PUBLIC ADMINISTRATION / ADMINISTRATION PUBLIQUE DU CANADA
VOLUME 58, NO. 1 (MARCH/MARS 2015), PP. 183–203

kept pace with these developments. Holding programs accountable for compliance with expectations seems clear and fair enough. But, by itself, this inevitably misleads and distorts. Such practices do not provide a proper account of the appropriateness and value of public sector activities and performance. Worse, they act as a disincentive to good performance and can lead to a wide range of perverse activities and outcomes. They provide the illusion rather than the reality of a focus on results and on accountability. This in turn contributes to the paradox that despite more resources and attention devoted to accountability-related activities, the public often feels that government is less and less accountable.

This article, after a brief consideration of what is meant by "accountability" and the nature and key implications of a results or outcome orientation, discusses shortcomings of traditional approaches to accountability. It considers reasons for the persistence of an approach to accountability with such shortcomings that have been so well documented, and then presents a vision of accountability more in keeping with current realities of public sector management.

The meaning of accountability in a results context

Gray and Jenkins (1993: 55) define accountability as: "the obligation to present an account of and answer for the execution of responsibilities to those who entrusted those responsibilities." Bemelmans-Videc (2007), the Canadian Comprehensive Audit Foundation (CCAF, in Leclerc et al. 1996) and others present similar definitions that, however, leave it open to whom and for what one is expected to be accountable, and how one appropriately can answer for a responsibility or power that has been conferred.

As Lonsdale and Bemelmans-Videc (2007: 3) observe, "Like honesty and clean water, 'accountability' is invariably seen as a 'good thing'." Nevertheless, Behn (2001) observes, in a book devoted specifically to the meaning of democratic accountability, that no one knows exactly what it means to "hold someone accountable" – except those who are being held accountable, where they understand that accountability in practical terms means punishment. Leclerc et al. (1996) and Thomas (2007) make similar observations, referring specifically to the Canadian context.

Lack of clarity about what is meant when referring to accountability limits the meaningfulness of the term. In practice, as Light (1993), Leclerc et al. (1996) and others have observed, accountability generally has been viewed as assessing compliance with tightly drawn rules and regulations. Such a traditional approach to accountability may well be appropriate to control against the abuse or misuse of power, to provide assurance that resources have been used for their intended purposes with due regard for fairness and good stewardship, attesting to the accuracy of accounts, and guarding against various forms of fraud and misrepresentation. However, it is less so with respect to performance.

But there is another face of accountability that has emerged with New Public Management (NPM) and now overshadows traditional accountability. The Auditor General of Canada (2002) has observed that one of the key purposes of accountability is to lead to improved performance of programs and policies. But as Dubnick (2005: 378) observes, the underlying assumption that "greater accountability will mean improved performance" remains largely unchallenged. As this article indicates, there is extensive evidence that traditional approaches to accountability are ill suited for the purpose of improving performance, leading to excessive focus on processes and on following procedures. Worse, this frequently results in unintended perverse effects, where accountability *practices* lose sight of the meaning of the concept.

Traditional approaches to accountability are ill suited for the purpose of improving performance.

The *concept* of accountability represents an important ethical and moral principle, basic to the concept and exercise of authority within a democracy (for example, Leclerc et al. 1996; Bemelmans-Videc 2007). A major purpose of accountability (some would say its prime function) is the legitimization of the exercise of authority, including the most appropriate use of public resources. In this sense, accountability can be viewed as an end in itself, with the objective of providing for greater confidence or assurance in what government is doing and how.

But from an outcome perspective, one must ask if accountability activities help *contribute* to more relevant, efficient, and effective public services. Looked at in this light, it is appropriate to ask if certain types of accountability approaches are more effective than others (for example, Jarvis and Thomas 2012; Perrin, Bemelmans-Videc and Lonsdale 2007). If accountability does not contribute in at least some way to improved government performance and effectiveness, then one may very well question its value. Worse is when questionable approaches to accountability negatively impact the legitimacy of our governance system.

A focus on results and what this means for public sector management and accountability

Arguably, the most important change affecting public sector management has been a shift from a primary focus on process or outputs to *a focus on* *outcomes and impacts*. Outcomes are fundamentally different in nature from processes, inputs, and outputs (or products). For example: they typically have a long-term trajectory, progress is not linear or incremental in nature and is likely to happen unpredictably in fits and starts with the likelihood of tipping points, and by their very nature may not be evident or measurable for some time. Typically there are numerous steps in the results chain, so that outcomes generally arise indirectly, and invariably not just as the result of a single (program) intervention, but in combination with other interventions, actions of other players, and mediated by social, economic and environmental factors.

There is no direct cause-and-effect relationship between activities and outcomes. Theories of change that take into account multiple actors, influencing factors, and feedback loops are required, in contrast to linear results chains that are inaccurate and can mislead (Rogers 2008). Rigid plans based upon pre-identified goals and targets (and logic models) are almost certain to become out of date, in response to changing priorities, needs, opportunities, threats, and feedback.

"Managing in a context of uncertainty" requires a flexible and adaptive style rather than a rigid approach based upon pre-identified outputs and targets (Freedman 2013; Handy 1995; Mintzberg 1994). Traditional approaches to accountability "customarily become associated with the judgment of whether a program or a policy has achieved its objectives," (Lehtonen 2005: 175) invariably through checking compliance against predetermined targets. But as the Canadian Comprehensive Audit Foundation has acknowledged, "rendering an account against an original plan without taking into account changed circumstances ... is not good accountability" (Leclerc et al. 1996: 34).

"Managing in a context of uncertainty" requires a flexible and adaptive style rather than a rigid approach based upon pre-identified outputs and targets.

Assessing performance and rewarding managers or programs for doing what they were expected to do, on the surface, sounds reasonable. But the reality is that this is inappropriate for an outcome focus when this involves assessing performance against pre-defined targets or indicators, as is the norm with most traditional accountability or results-based management (RBM) approaches. As Mayne and Zapico-Goñi (1997) observed: "With an increasingly diverse, interdependent, and uncertain public sector environment, ... meeting objectives fixed some time ago may not be as important as the capacity to adapt to current and future change... In a rapidly changing world, responsive programs *should* be changing and modifying, and should be rewarded rather than penalized for doing so."

Implementing an outcome focus requires a significant change in the ways of thinking and approach to management concerning *all* aspects of government, including reward mechanisms and accountability approaches. A roundtable discussion convened by the World Bank (Perrin 2006), involving international experts considered what is needed for governments to move from a focus on outputs to outcomes. While affirming that such a strategic focus is central to the *raison d'être* of government, the experts highlighted the challenges of such a change, indicating that one should not expect perfection, and to encourage trying, should reward "failure," provided that learning comes from this. The Auditor General of Canada (2002: 3) has observed: "Public sector management and governance are changing, becoming more complex and creating new pressures on traditional notions of accountability. Thus it should hardly be surprising that this has not proved easy to implement."

An expert meeting OECD convened specifically to discuss challenges to results-focused management and budgeting provides a good illustration of challenges to an results-oriented approach, and how it can be sabotaged by inappropriate measures more suited for keeping track of activities. Despite the stated objective and title of the event, many of the countries represented said that outcomes were too difficult to measure and in particular to be useful for accountability because of difficulties of attribution. As a result, there was minimal attention to outcomes, some to outputs, but with primarily attention still placed on the development of better means of controlling inputs (Perrin 2002). One may well ask (for example, as Ohemeng and McCall-Thomas 2013 and Radin 2011, among others, have done) if the "results orientation" is more rhetoric than reality.

Connecting accountability to performance measurement

As many have indicated (for example, Perrin 1998, 1999, 2002, 2012; Hatry 2013; Hildebrand and McDavid 2011; McDavid, Huse and Hawthorn 2013; Nielsen and Hunter 2013) performance measures can be a useful tool for management in many circumstances. In particular, performance measures need to be viewed as *indicators*, as one element of a more comprehensive monitoring and evaluation strategy, for *raising* rather than answering questions that could then be explored through means ranging from some telephone calls to a comprehensive evaluation. The main problem occurs when *consequences* are placed upon achieving pre-established targets, such as for external reporting, rewards and punishments, and accountability, leading to distortion and misrepresentation of actual performance.

There are numerous reasons why it is inappropriate to use indicators and targets for accountability purposes. Their use is inconsistent with evidence that it is not appropriate to reduce complex undertakings, representing almost all significant public sector initiatives, to one or a small number of (primarily) quantitative indicators (for example, Forss, Marra and Schwartz 2011; Freedman 2013; Handy 1995; Hummelbrunner and Jones 2013a; Mintzberg 1994, 1996; Newcomer 1997; Rogers 2008; Williams and Hummelbrunner 2011). Even more basically, reliance on indicators assumes that they present a valid picture of what they purport to measure. But as Dubnick 2005, Tsoukas 2004 and others have indicated, this is highly questionable on epistemological grounds. Choosing indicators is invariably based upon negotiation, bureaucratic and political pressures, and thus in spite of their apparent objectivity, they are inherently selective, subjective and value laden. At best, indicators present a narrow window on reality.

Traditional accountability approaches may provide a misleading account of the appropriateness and effectiveness of performance, and worse, leads to perverse incentives and to the undermining of effectiveness.

Inappropriate quantification and a high-stakes approach to accountability often means that attention is paid to what is easy to measure, to inputs, activities and, perhaps, outputs, rather than to outcomes. This can sabotage a meaningful outcome focus. Consequently, traditional accountability approaches may provide a misleading account of the appropriateness and effectiveness of performance, and worse, leads to perverse incentives and to the undermining of effectiveness. Even when intended as a management tool, RBM approaches frequently become *de facto* accountability and control mechanisms, or at least are frequently viewed and used in this manner (for example, Gill 2011; Mintzberg 1996).

Given these shortcomings, it also is hardly surprising that traditional approaches to accountability, based upon performance measures, fail in their objective of providing for greater confidence in what government is doing and benefits arising.

Performance measures can fail to provide a meaningful account of actual performance

There is extensive evidence documenting the limitations of performance measures to present an accurate and meaningful view of actual performance, thereby misrepresenting what programs are actually doing.[1] It is only possible to discuss a few of these factors below (but see, for example, Perrin 1998, for discussion of other barriers and considerations that limit the ability of performance to give a valid and accurate accounting of actual performance).

Meaningless and inaccurate data

Transocean, the deep-water drilling company responsible in part for the explosion of the BP Deepwater Horizon oil rig in the Gulf of Mexico that killed a number of people and resulted in a major ecological disaster nevertheless awarded its executives substantial bonuses, claiming in its annual report that:

> "Notwithstanding the tragic loss of life in the Gulf of Mexico, we achieved an exemplary *statistical* safety record as measured by our total recordable incident rate and total potential severity rate . . . As measured by these standards, we recorded the best year in safety performance in our company's history" (cited by BBC News 2011, emphasis added).

This represents just one poignant example, among many that are well documented in the literature, of how performance indicators can be meaningless. As another example, the European Union Cohesion (regional development) Policy uses numbers of people "employed" and "in training" to hold member states to account. But what do these terms mean? Numbers of people "in employment" can include people in "good"' well-paying steady full-time jobs with a future, as well as others in part-time work of a few days' duration in an abusive situation; "self employment" (which is included as a form of employment) can include street hawkers with insufficient revenue to cover their expenses as well as others running major hi-tech enterprises with a high level of income. Numbers of people "trained" can include those who slept through a half-day session of dubious value (or count the same person multiple times for registering for many such sessions) as well as those who have completed a comprehensive six-month training program (Perrin 2011).

There are many cases where definitions of "obvious" measures vary tremendously (for example, "client," which can, and is, defined in numerous ways, even among agencies providing similar services, and then inappropriately aggregated). Often clerical staff required to submit data, along with busy professionals who resent the time they need to devote to "paperwork" rather than to delivering services, make it up as they go along.

Goal displacement

When indicators become the objective, they can result in "goal displacement" which leads to emphasis on "making the numbers" rather than on doing what the program was supposed to be doing. At best, this can result in performance measures misrepresenting actual performance.

Examples of goal displacement are well documented. For example, the UK Work Programme uses a payment-by-results mechanism whereby private sector contractors are paid only on the successful placement of welfare recipients into employment. An independent evaluation (Rees, Taylor and Damm 2013) not only found little success, but that "creaming" and "parking" were embedded in the approach, where those closest to the labour market received the most help and those with the greatest need received minimal attention. In an empirical study concerning standardized testing in Ontario public schools, Ohemeng and McCall-Thomas (2013) document how centralized imposed targets for "results" have led to undesired behaviours such as "the opening of the test before test day, giving students the questions before the test, erasing test answers after the exams, teaching to the test, and circumventing the instructions for test administration" (466).

There are numerous other ways of meeting targets without improving performance, such as: selective definitions, for example, relabeling dropouts from a program as having moved or even as a "success" in becoming independent; "encouraging" those who do not seem to be doing well to go elsewhere in order to increase the overall success rate; providing a minimal level of service when the target is numbers of people served; hospitals addressing waiting time standards by providing perfunctory attention and thereby stopping the clock while keeping patients waiting for hours to receive any real treatment, ignoring those who are not going to meet a target, such as call centres that terminate calls that cannot be answered in time. Ways of gaming the system in order to meet targets without addressing real needs are widespread and well documented (for example, Perrin 1998; Thomas 2006; Wheelan 2013).

Incentive to distort results

Holding managers accountable to meet targets can produce strong pressure to misrepresent and to distort results, with potentially devastating consequences. It is now apparent that many of the recent corporate scandals are at least partially the result of inappropriate incentives and inordinate pressure on managers to "meet their numbers" — indeed all too often to "meet their *number*." For example, Enron indicated in its annual report that it was "laser-focused on earnings per share," and its former chief executive, Kenneth Lay, reportedly received a bonus of US$123M for achieving his target, at the same time that the company was virtually defunct.

This is hardly an isolated incident. For example, an article in the *Harvard Business Review* indicates that the use of targets to determine compensation "encourages managers to lie and cheat, lowballing targets and inflating results, and it penalizes them for telling the truth. It turns business decisions into elaborate exercises in gaming. It sets colleague against colleague, creating distrust and ill will. And it distorts incentives, motivating people to act in ways that run counter to the best interests of their companies." (Jensen 2001: 94).

The same sort of thing also happens in the public sector. Linking pay to target achievement is happening more and more, in spite of evidence that pay-for-performance does not work (for example, Bevan 2013; Chartered Institute of Personnel and Development [CIPD] 2009; Toynbee 2013). Threatening to "punish failure" puts a high degree of pressure on even public sector managers to meet their targets, no matter how, if they want to keep their jobs and to preserve their programs. Bevan and Hamblin, citing a UK Audit Commission report, indicate that: "In a culture where managers' jobs depend on achieving specific targets, there will be pressure to meet those targets (2009:182). Studies in jurisdictions such as France (L'Inspection Générale de l'Administration et Inspection Générale de la Police Nationales 2014) and the UK (House of Commons 2014) have documented how police have intentionally manipulated statistics on crime in order for "crime" to more closely match political expectations.

Critical subgroup differences hidden

As Hatry (2013: 25) has observed: "Performance measurement systems typically provide only aggregate outcome data." However, aggregate (average) scores can misrepresent what is really happening. For example, Winston (1993) indicates how a program with 40 percent women participants can achieve a 60 percent "success" rating, with all – or none – of the women being successful. This finding would be hidden by an indicator that only looked at the overall success rate.

Similarly, average scores showing overall improvement in household income may disguise the fact that while a few people (typically those in the top decile) are doing better, most people are no better off, with those at the bottom end actually worse off. As Pawson and Tilley (1997) have observed, with almost any program, invariably some people are better, and others worse off. Performance measurement data that fail to show this are at best misleading.

Inhibits rather than contributes to improved performance

The above discussion indicates how performance measures may fail to provide a meaningful perspective on actual performance, and thus undermine accountability or decision making-related uses. But more than this, there is also extensive evidence about how traditional accountability approaches combined with results-based schemes based upon indicators adversely affect performance.[2]

Silo thinking and action

Few significant objectives (for example, poverty reduction, economic development, employment creation, crime reduction, health, climate change) can be addressed without coordinated action across multiple program areas or without involvement of non-governmental partners. Outcomes also depend upon factors beyond the direct control of a single manager or program.

> *Traditional accountability approaches combined with results-based schemes based upon indicators adversely affect performance.*

Nevertheless, most reward and accountability structures remain vertical. More appropriate shared accountability mechanisms are rare. This can serve as a powerful disincentive to cooperative action, and may result in cost shifting to other areas and to inappropriate attention to processes and outputs more within the control of a manager.

Diversion of limited resources

Accountability does not come cheaply (for example, Power 1997; Gray and Jenkins 2007; Martin 2005). Resources that otherwise could go toward program improvement or delivering services must instead be devoted to documentation, a major complaint of program staff. As Gray and Jenkins (2007) ask, what should be the appropriate balance between checking and doing?

Programs charged with delivering services are expected (often repeatedly) to demonstrate their value and cost effectiveness. Surely the same should apply to accountability-related functions (for example, RBM specialists, audit, inspection, evaluation, other forms of monitoring) that are of value only if they can assist in improving the relevance, effectiveness, and efficiency of public services. It is incumbent upon all those engaged in accountability to pay attention to whether they are delivering benefits otherwise they are part of the problem rather than part of the solution.

Less, rather than more, focus on results and on innovation

The sad, supreme irony is that "results-oriented" accountability approaches typically lead to less – rather than to more – focus on outcomes, on innovation and improvement. As Gill (2011) and Mintzberg (1996) have indicated, this approach is rooted in a top-down hierarchical "control" model. A narrow focus on measurement is inconsistent with a focus on change and improvement that requires constant questioning about what else can be done or done better (Senge 1990). The outcome of judging programs or staff by their "numbers" is not empowerment, critical self-examination, and experimentation with new approaches. Instead, it leads to impaired performance, an emphasis on justifying and defending what was done, and a reluctance to admit that improvement is needed.

A compliance approach may indeed be appropriate with respect to accountability for proper use of funds, conformity with legal or safety requirements, and perhaps with other matters that can be clearly specified. But a different approach is required regarding accountability for results.

Meaningless for decision making

Perhaps the most frequently mentioned rationale for performance measurement, including (indeed, often particularly) for accountability purposes, is to provide for more informed decision making and budgeting (for example, Hatry 1997). But performance measures, by themselves, are useless for this purpose. As Newcomer (1997: 10) put it: "Performance measurement typically captures quantitative indicators that tell what is occurring with regard to program outputs and perhaps outcomes but, in itself, will not address the how and why questions." Programs may fail to meet targets due to limitations of program design – but also due to faulty management or implementation, under (or over) funding, unique circumstances, inappropriate targets or measurement at the wrong time, or for other reasons (Perrin 1998).

Using performance measures ... for accountability purposes, including making decisions about the future of programs without other forms of supporting evidence, is likely to lead to inappropriate actions.

Thus performance measures by themselves provide no direct implications for action, unless other means are used to explore the reasons for results and the potential for future impact. Indeed, using them for accountability purposes, including making decisions about the future of programs without other forms of supporting evidence, is likely to lead to inappropriate actions.

Less rather than more confidence in government

Lonsdale and Bemelmans-Videc (2007: 3) have indicated that in spite of the "paradox of there being more accountability-related activities today than ever before, at the same time ... much public debate laments what is seen as a lack of actual accountability." Tsoukas (2004) identifies paradoxes of the information society, where more information such as performance indicators devoid of context results in less understanding and in particular tends to erode trust that raises legitimate questions about what is being concealed. Radin observes that: "Citizens across the globe are skeptical about the ability of their governments to be accountable or able to perform as expected" (2011: 98). A study concerning accountability for funding for child-care services in Canada found lack of interest and trust in public reporting, with widespread distrust about the veracity of government figures (Anderson and Findlay 2010).

While there are many factors at play, current accountability approaches do little to address concerns of the public about the effective use or misuse of government resources. This is hardly surprising given the failure of accountability approaches to reflect actual performance and meaningful results, what is seen as excessive and inappropriate bureaucracy and sophistry with statistics, and failure to support improved performance.

Why such persistence with an inappropriate approach to accountability?

There has been extensive documentation over the years of the limitations of traditional approaches to accountability and RBM – within Canada as well as in numerous other jurisdictions, including several articles in recent issues of this journal (for example, Anderson and Findlay 2010; Hildebrand and McDavid 2011; Ohemeng and McCall-Thomas 2013; Thomas 2007). These limitations and distortions are, or at least should be, well known. Nevertheless, there is an apparent reluctance to consider anything different and indeed a proliferation of RBM approaches and accountability measures that reflect many of the difficulties with these approaches. Why is this?

Do governments, or at least politicians and high-ranking officials, ***really*** *want to improve?* Perhaps politicians and other senior officials are not really concerned about the reality of government performance, but primarily want a means of making themselves look good (and a means of blaming others where there are problems) and a way of symbolically suggesting that they are concerned about accountability and results? Perhaps the "results agenda" is more a matter of rhetoric and ideology than reality and rational analysis?

This is, arguably, a cynical perspective. But it is one frequently suggested by observers (for example, Radin 2011), and it is one that I hear frequently, in many different jurisdictions, from people both inside and outside of government. One can, perhaps, understand this attitude among politicians and some managers who might feel threatened by a meaningful results orientation. But I am convinced that most public officials are in public service because they *are* concerned about doing the best for their constituency. And is there not an obligation for auditors, evaluators, and others engaged in the accountability enterprise to speak truth to power, rather than exacerbating the problem?

Control trumps good management. One reason for accountability is to control. Government officials want control over what is done with "their" funds and how the direction they have mandated is implemented. The problem – or at least one of the problems with this approach – is that "control" approaches simply do not work, at least with respect to bringing out the best in managers and staff.

The control or machine model of human resources management went out of fashion some time ago – because it is ineffective. A top-down command-and-control approach to management is not suitable for creating a meaningful results focus (for example, see Mintzberg 1996). As a report of an expert OECD meeting considering challenges to results-focused management and budgeting observed: "One can order people to undertake specific activities. But it is impossible to order or to direct people how to think or what to believe. Indeed, this is most likely to be counterproductive" (Perrin 2002:14). There are much more effective ways of getting others to do what one wants.[3]

Leadership from the top is needed to bring about and to support needed organizational renewal and change.

So, why the persistence with control mechanisms that impede performance? Is the appearance of exerting power and the illusion of control more important than what actually results from it? As Freedman (2013:557) states: "Power certainly could become an end in itself, a source of status and opportunities to boss others around [that is] detrimental to overall efficiency as well as to morale."

Demand for simple (or simplistic) answers to complex policy problems. There often is a demand for simplicity, in particular by politicians, the media, and the public. Thirty-second sound bites have become the medium for communicating governance issues. Explaining a complex issue through just a few indicators confuses simplicity and simplistic communications.

In part, this demand follows from an inappropriate obsession with numbers and quantification, a simplistic belief that numbers are never wrong.

Unaware of alternatives. Even those well aware of the limitations of indicators sometimes do not seem to be able to conceive of any other approach, and instead seem to call for more of the same, but just done better, in spite of the substantial evidence that indicators, by themselves, are almost certain to be misused, in both the private (for example, The Economist 2002; Handy 1995, Jensen 2001; Wheelan 2013) and public sectors (for example, Bevan and Hood 2006; Le Grand 2010; Perrin 2008). There seems to be lack of openness to consideration of means of assessing performance and of accountability other than through quantitative indicators, despite amble evidence of their shortcomings.

Resistance to change. Inertia is a powerful force. Why change unless one is really forced to do so? And there are many with an entrenched interest in staying with the status quo, which is generally easier – at least in the short run.

A shift to a true outcome orientation, including realignment of accountability approaches, represents a profound paradigm shift and a major change in mindset. The implications of moving to a focus on outcomes rather than on process are still not understood, or accepted.

A vision of accountability appropriate to undertakings intended to achieve outcomes that take place in a complex policy environment and are inevitably influenced by a variety of factors.

Leadership from the top is needed to bring about and to support needed organizational renewal and change. There is still limited, but an increasing array of resources available about how to manage for outcomes in a way that embraces complexity.[4] Organizations that remain static and fail to evolve and improve quickly become out of date and may struggle to survive, at least in the long term.

A better approach to accountability

Perrin, Bemelmans-Videc and Lonsdale (2007) present a vision of accountability appropriate to undertakings intended to achieve outcomes that take place in a complex policy environment and are inevitably influenced by a variety of factors.

This approach to accountability encompasses three essential characteristics:

- A primary orientation toward results rather than on process.
- A focus on continuous and responsive learning.

- A dynamic rather than a static approach that reflects the complexities and uncertainties inherent in most public policy areas.

This model of accountability involves holding programs accountable for: asking the tough questions about what works and why, innovating and engaging in risk taking rather than playing it safe, and for seeking – and using – feedback. Holding programs accountable for asking the difficult questions, doing and using evaluations, and demonstrating use of learning – such as through changes in policies and in program approaches, may represent a harder standard than demonstrating compliance with procedures as with traditional accountability.

This approach to accountability is based upon the following principles:

- Acting *responsibly* – being trustworthy, being true to the mandate, demonstrating responsibility in taking decisions (for example, Gregory 1995).
- Addressing the overall need or rationale for why the program is in place.
- Doing the best possible job given the circumstances, resources and constraints, consistent with the overall mandate.(for example, Light 1993).

In short, programs should be accountable for demonstrating good management and for keeping in view outcomes, which includes (but definitely is not limited to) a true results orientation.

A results-oriented focus represents a new way of thinking and managing, that needs to be reflected in all *aspects of management.*

Such an approach to accountability requires a change in the processes that characterize traditional compliance-oriented accountability approaches. This approach is consistent with views advocated by many respected authorities. For example, it is consistent with the approach advocated by the Auditor General of Canada (2002: 10):

> "Our enhanced concept of accountability supports managing for results and hence a culture of learning. It asks ministers and managers to demonstrate credibly that they are learning (from mistakes as well as successes), taking corrective action where appropriate, and following up on weaknesses, rather than focussing only on who is at fault when things go wrong."

The Auditor General further says that holding to account for results asks if "everything reasonable has been done with available authorities and resources to influence the achievement of expected results" (2002: 8) and that while people should accept responsibility for their mistakes:

"If we wish to empower employees and encourage them to innovate ... we should focus on learning from the experience rather than assigning blame" (2002: 10). Mayne (2007) has further elaborated upon these points.

The World Bank roundtable discussion of experts identified a number of principles for moving towards a true results-oriented approach:

- Recognition that a results-oriented focus represents a new way of thinking and managing, that needs to be reflected in *all* aspects of management. Need for strong demonstrated, tangible and visible commitment from the top political and administrative levels, in order to provide legitimacy and priority to an outcome orientation and mobilization of resources as required. Need for *bottom-up* support and engagement, otherwise an outcome focus runs the risk of becoming a mere administrative exercise rather than an actual change in approach.
- Need to support and reward innovation and risk taking, being careful not to punish those who try, even if initial efforts are not perfect.
- Need to be strategic, relating *all* aspects of the results-oriented approach to the strategic direction and goals.
- Monitoring and evaluation approaches should not be developed in a vacuum, but in response to information requirements that will be needed to inform decisions and future directions.

One might take the approach developed initially within the Canadian Office of the Auditor General (Mayne 2001, 2011), where indicators, along with other sources of information, are used to develop the performance story, or as Winston (1999) has put it, one assembles a variety of forms of data in order to provide performance *information* that is balanced between being formative (intended to improve existing policies and programs) and summative (intended to publicly sum up achievements and shortcomings).

Performance information regarding what is really important cannot be reduced to a few numbers in a database.

A corollary of the above is that there is a need for program evaluation as well as for monitoring (or RBM or performance management) as part of a more comprehensive monitoring and evaluation strategy to provide a meaningful picture of how effectively programs are moving towards outcomes, and in particular to provide explanation that can better inform improvements as well as future policies and strategies. It makes little sense, as is too frequently the case, to have RBM and evaluation functions completely separate (e.g. see Newcomer and Brass 2015 or Nielsen and Hunter, 2013: a special issue of *New Directions for Evaluation* devoted specifically to this topic).

Conclusion and next steps

There is no question that implementing the above vision represents a hard sell. In particular, it requires recognition that performance information regarding what is really important cannot be reduced to a few numbers in a database. This may represent a different view of the world for many of those whose primary training and expertise has been with such means and where "accountability", as Behn (2001) has stated, is often equated with punishment.

Given the above, how can one get leadership buy-in for a more meaningful approach to accountability in cultures where some say that the adversarial nature of politics and any moves to decouple performance from a hard-edged view are viewed, at best, with scepticism? It is well beyond the scope of this article to discuss this in any detail, and I can only offer a couple of ideas.

If reforms are not undertaken, "accountability" practices will continue to be more and more rhetorical, inhibiting improved performance and contributing to less rather than more confidence in government.

First, there is a need for greater public acknowledgement of the widely known shortcomings of current approaches to accountability, and then discussion about possible alternatives. Perhaps this journal and/or others with an interest in the area (for example, the Institute of Public Administration of Canada, Canada's Treasury Board or the Office of the Auditor General) could provide a forum where leading thinkers, including public sector leaders, can discuss such considerations. Internationally, a Global Parliamentarian Forum on Evaluation is being launched during 2015. This might provide an opportunity for those parliamentarians interested in outcome-oriented and effective public services to discuss with their peers possible roles for parliamentarians in supporting such a change.

Moving to a model of accountability appropriate for the current realities of public governance represents a major change effort. Nevertheless, if reforms are not undertaken, "accountability" practices will continue to be more and more rhetorical, inhibiting improved performance and contributing to less rather than more confidence in government.

Notes

1 For example, Bemelmans-Videc et al. 2007; Behn 2001; CIPD 2009; de Lancer Julnes 2006; Feller 2002; Hummelbrunner and Jones 2013a, 2013b; McDavid and Huse 2012; McDavid, Huse, and Hawthorn 2013; Nielsen and Hunter 2013; Perrin 1998, 2002; Pollitt 2013; Power 1997; Thomas 2006; van der Knapp 2006; Winston 1999.
2 For example, see Note 1.
3 Discussion of this is well beyond the scope of this article, but see, for example: Behn 2004; CIPM 2009; Mintzberg 1996; Hummelbrunner and Jones 2013a, 2013b; Perrin 2002, 2006. Or almost any human resources text.
4 To give but some examples: Bemelmans-Videc et al. 2007; Behn 2001; Forss, Marra and Schwartz 2011; Perrin 2002, 2007; Williams and Hummelbrunner 2011; Williams and Imam 2007.

References

Anderson, Lynell, and Tammy Findlay. 2010. "Does public reporting measure up? Federalism, accountability and child-care policy in Canada." *Canadian Public Administration*, 53 (3): 417–438.

Auditor General of Canada. 2002. *Modernizing Accountability in the Public Sector*. Chapter 9, Report of the Auditor General of Canada to the House of Commons. Ottawa.

BBC News. 2011. "Transocean gives bonuses after Gulf of Mexico BP spill." 3 April.

Behn, Robert D. 2001. *Rethinking Democratic Accountability*. Washington: Brookings.

——. 2004. "Performance leadership: 11 better practices that can ratchet up performance." IBM Center for the Business of Government. Retreived May. [website] Available at: http://www.businessofgovernment.org/pdfs/Behn_Report.pdf.

Bemelmans-Videc, Marie-Louise. 2007. "Accountability, a classic concept in modern contexts: Implications for evaluation and for auditing roles." In *Making Accountability Work: Dilemmas for Evaluation and for Audit*, edited by Marie-Louise Bemelmans-Videc, Jeremy Lonsdale, and Burt Perrin. New Brunswick, NJ: Transaction Publishing.

Bemelmans-Videc, Marie-Louise, Jeremy Lonsdale, and Burt Perrin, eds. 2007. *Making Accountability Work: Dilemmas for Evaluation and for Audit*. New Brunswick, NJ: Transaction Publishing.

Bevan, Steven. 2013. "Performance pay won't perform in the classroom." The Work Foundation Blog. Posted 22 October. Available at: http://www.theworkfoundation.com/blog/1438/Performance-pay-wont-perform-in-the-classroom#.UoH7SJBk0hQ.email.

Bevan, Gwen, and Christopher Hood. 2006. "What's measured is what matters: targets and gaming in the English public health care system." *Public Administration* 84 (3): 517–538.

Bevan, Gwen, and Richard Hamblin. 2009. "Hitting and missing targets by ambulance services for emergency calls: Effects of different systems of performance measurement within the UK." *Journal of the Royal Statistical Society* 172, Part 1: 161–190.

Chartered Institute of Personnel and Development (CIPD). 2009. *Performance Management in Action: Current Trends and Practice*. London: author.

de Lancer Julnes, Patria. 2006. "Performance measurement: An effective tool for government accountability? The debate goes on." *Evaluation* 12 (2): 219–235.

Dubnick, Melvin. 2005. *Public Performance & Management Review* 28 (3) March: 376–417.

Economist, The. 2002. "Why honesty is the best policy: Corporate deceit is a slippery slope." *The Economist* 362 (8263), March.

Feller, Irwin. 2002. "Performance measurement redux." *American Journal of Evaluation* 23 (4): 435–452.

Forss, Kim, Mita Marra, and Robert Schwartz, eds. 2011. *Evaluating the Complex: Attribution, Contribution, and Beyond*. New Brunswick, NJ: Transaction Publishers.

Freedman, Laurence. 2013. *Strategy: A History*. New York: Oxford University Press.
Gill, Derek, ed. 2011. *The Iron Cage Recreated: The Performance Management of State Organizations in New Zealand*. Wellington: Institute of Policy Studies, Victoria University.
Gray, Andrew, and Bill Jenkins. 1993. "Codes of accountability and in the new public sector." *Accounting, Auditing, and Accountability Journal*, 6 (3): 52–67.
——. 2007. "Checking out? Accountability and evaluation in the British regulatory state. In *Making Accountability Work: Dilemmas for Evaluation and for Audit*, edited by Marie-Louise Bemelmans-Videc, Jeremy Lonsdale, and Burt Perrin. New Brunswick, NJ: Transaction Publishing.
Gregory, Robert. 1995. "Accountability, responsibility and corruption: Managing the public production process." In *The State Under Control*, edited by Jonathan Boston. Wellington, NZ: Bridget Williams Books.
Handy, Charles B. 1995. *Beyond Certainty: The Changing Worlds of Organizations*. Boston: Harvard Business School Press.
Hatry, Harry P. 1997. "Where the rubber meets the road: Performance measurement for state and local public agencies." In *Using Performance Measurement to Improve Public and Non-profit Programs*. New Directions for Evaluation, No. 75, edited by Kathryn E. Newcomer San Francisco: Jossey-Bass.
Hatry, Harry P. 2013. "Sorting the relationship among performance measurement, program evaluation, and performance management." In *Performance Management and Evaluation*, edited by Steffen Bohm Nielsen, and David E. K. Hunter. New Directions for Evaluation No. 137. San Francisco: Jossey-Bass.
Hildebrand, Richard, and James C. McDavid. 2011. "Joining public accountability and performance management: A case study of Lethbridge, Alberta." *Canadian Public Administration* 54 (1): 41–72.
House of Commons (UK) Public Administration Select Committee (PASC). 2014. *Caught Red-Handed: Why We can't Count on Police Recorded Crime Statistics*. Thirteenth Report of Session 2013–14. London: The Stationery Office Limited. Also available at http://www.publications.parliament.uk/pa/cm201314/cmselect/cmpubadm/760/760.pdf.
Hummelbrunner, Richard, and Jones, Harry. 2013a. *A Guide to Managing in the Face of Complexity. Background Note*. London: Overseas Development Institute. http://www.odi.org.uk/sites/odi.org.uk/files/odi-assets/publications-opinion-files/8662.pdf.
——. 2013b. *A Guide for Planning and Strategy Development in The Face of Complexity. Background Note*. London: Overseas Development Institute. http://www.odi.org.uk/sites/odi.org.uk/files/odi-assets/publications-opinion-files/8287.pdf.
L'Inspection Générale de l'Administration et Inspection Générale de la Police Nationales. 2014. *L'Enregistrement des Plaintes par les Forces de Sécurité Intérieure sur le Ressort de la Préfecture de Police*. N° 14-011/13-093/01 et IGPN-SG-N° 13-1000 I. Paris: Ministère de l'Intérieur. Available at http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/144000135/0000.pdf.
Jarvis, Mark D., and Paul Thomas. 2012. "The Limits of Accountability: What Can and Cannot Be Accomplished in the Dialectics of Accountability?" In *From New Public Management to New Political Governance: Essays in Honour of Peter C. Aucoin*, edited by Herman Bavkis and Mark D. Jarvis. Montreal: McGill-Queen's University Press.
Leclerc, Guy, David Moynagh, Jean-Pierre Boisclair, and High R. Hanson. 1996. *Accountability, Performance Reporting, Comprehensive Audit – An Integrated Perspective*. Ottawa: Canadian Comprehensive Audit Foundation (CCAF).
Le Grand, Julian. 2010. "Knights and knaves return: Public service motivation and the delivery of public services." *International Public Management Journal*, 13 (1): 56–71.
Lehtonen, Markku. 2005. "OECD environmental performance review programme: Accountability (f)or learning?" *Evaluation* 11 (2): 169–188.

Light, Paul C. 1993. *Monitoring Government: Inspectors General and the Search for Accountability*. Washington: Brookings Institution Press.

Lonsdale, Jeremy, and Marie-Louise Bemelmans-Videc. 2007. In *Making Accountability Work: Dilemmas for Evaluation and for Audit*, edited by Marie-Louise Bemelmans-Videc, Jeremy Lonsdale, and Burt Perrin. New Brunswick, NJ: Transaction Publishing.

Martin, Steve. 2005. "Evaluation, Inspection and the Improvement Agenda: Contrasting Fortunes in an Era of Evidence Based Policy Making." *Evaluation* 11 (4): 496–504.

Mayne, John. 2001. "Addressing attribution through contribution analysis: Using performance measures sensibly." *Canadian Journal of Program Evaluation* 16 (1): 1–24.

——. 2007. Evaluation for accountability: Myth or reality? In *Making Accountability Work: Dilemmas for Evaluation and for Audit*, edited by Marie-Louise Bemelmans-Videc, Jeremy Lonsdale, and Burt Perrin. New Brunswick, NJ: Transaction Publishing.

——. 2011. "Contribution analysis: Addressing cause and effect." In *Evaluating the Complex: Attribution, Contribution, and Beyond*, edited by Kim Forss, Mita Marra, and Robert Schwartz, eds. New Brunswick, NJ: Transaction Publishers.

Mayne, John, and Eduardo Zapico-Goñi, eds. 1997. *Monitoring Performance in the Public Sector: Future Directions from International Experience*. Transaction Publishers.

McDavid, James. C., and Irene Huse. 2012. "Legislator uses of public performance reports: Findings from a five-year study." *American Journal of Evaluation* 33 (1): 7–25.

McDavid, James. C., Irene Huse, and Laura R.L. Hawthorn. 2013. *Program Evaluation and Performance Measurement An Introduction to Practice*. Sage: Thousand Oaks, CA.

Mintzberg, Henry 1994. *The Rise and Fall of Strategic Planning*. New York: The Free Press.

——. 1996. "Managing government, governing management." *Harvard Business Review* 75–83.

Newcomer, Kathryn E. 1997. "Using performance measurement to improve public and non-profit programs." In *Using Performance Measurement to Improve Public and Nonprofit Programs* edited by Kathryn E. Newcomer. *New Directions for Evaluation* No. 75. San Francisco: Jossey-Bass.

Newcomer, Kathryn E., and C. Brass. 2015. "Forging a strategic and comprehensive approach to evaluation within public and nonprofit organizations: Integrating measurement and analytics within evaluation" *American Journal of Evaluation*. In press.

Nielsen, Steffen Bohm and David E. K. Hunter, eds. 2013. *Performance Management and Evaluation*. New Directions for Evaluation, No. 137. San Francisco: Jossey-Bass.

Ohemeng, Frank, and Elyse McCall-Thomas. 2013. "Performance management and 'undesirable' organizational behaviour: Standardized testing in Ontario schools." *Canadian Public Administration* 56 (3): 456–477.

Pawson, Ray, and Nick Tilley. 1997. *Realist Evaluation*. London: Sage Publications.

Perrin, Burt. 1998. "Effective use and misuse of performance measurement." *American Journal of Evaluation* 19 (3): 367–379.

——. 1999. "Performance measurement: Does the reality match the rhetoric? A rejoinder to Bernstein and Winston." *American Journal of Evaluation* 20 (1): 101–110.

——. 2002. "Implementing the vision: Addressing challenges to results-focused management and budgeting." OECD. Available at: http://www.oecd.org/dataoecd/4/10/2497163.pdf.

——. 2006. *Moving from Outputs to Outcomes: Practical Advice from Governments Around the World*. Washington: World Bank and IBM Center for the Business of Government. Available at: http://www.worldbank.org/oed/outcomesroundtable/outputs_to_outcomes.pdf.

——. 2007. "Towards a new view of accountability." In *Making Accountability Work: Dilemmas for Evaluation and for Audit*, edited by Marie-Louise Bemelmans-Videc, Jeremy Lonsdale, and Burt Perrin. New Brunswick, NJ: Transaction Publishing.

——. 2011. "What is a true results-performance-based delivery system." Invited expert presentation to the European Parliament. *Evaluation* 17 (4): 17–424.

——. 2012. *Linking Monitoring and Evaluation to Impact Evaluation*. Guidance Note No.2 Washington: InterAction. Available at: http:www.interaction.org/impact-evaluation-notes

Perrin, Burt, Marie-Louise Bemelmans-Videc, and Jeremy Lonsdale. 2007. "How evaluation and auditing can help bring accountability into the twenty-first century." In *Making Accountability Work: Dilemmas for Evaluation and for Audit*, edited by Marie-Louise Bemelmans-Videc, Jeremy Lonsdale, and Burt Perrin. New Brunswick, NJ: Transaction Publishing.

Pollitt. Christopher. 2013. "The logics of performance management." *Evaluation* 19: 346–363.

Power, Michael. 1997. *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.

Radin, Beryl A. 2011. "Does performance measurement actually improve accountability?" In *Accountable Governance: Problems and Promises*, edited by Melvin J. Dubnick and H. George, Frederickson. Armonk, NY: M.E. Sharpe, 98–110.

Rees, James, Rebecca Taylor, and Chris Damm. 2013. *Does Sector Matter? Understanding the Experiences of Providers in the Work Programme*. Third Sector Research Centre Working Paper 92. Available at: http://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/working-paper-92.pdf

Rogers, Patricia. 2008. "Using programme theory to evaluate complicated and complex aspects of interventions." *Evaluation* 14(1): 29–48.

Senge, Peter. 1990. *The Fifth Discipline: The Art and Practice of the Learning Organization*. New York: Doubleday/Currency.

Thomas, Paul G. 2006. *Performance Measurement, Reporting, Obstacles and Accountability: Recent Trends and Future Directions*. Canberra: The Australian National University.

——. 2007. "Reflections of a 90s-era editor (the 1990s, that is, not the 1890s!)." *Canadian Public Administration* 50: 513–516.

Toynbee, Polly. 2013. "The push for performance-related pay is driven by faith, not facts." *The Guardian*, 12 November. Accessed at: http://www.theguardian.com/commentisfree/2013/nov/12/performance-related-pay-risky-behaviour.

Tsoukas, Haridimos. 2004. "The tyranny of light." In *Complex Knowledge: Studies in Organizational Epistemology*. Oxford: Oxford University Press.

van der Knaap, Peter. 2006. "Responsive evaluation and performance management: Overcoming the downsides of policy objectives and performance indicators." *Evaluation* 12 (3).

Wheelan, Charles. 2013. *Naked Statistics*. New York, Norton.

Williams, Bob, and Iraj Imam, eds. 2007. *Systems Concepts in Evaluation: An Anthology*. Point Reyes, CA: EdgePress of Inverness.

Williams, Bob, and Richard Hummelbrunner. 2011. *Systems Concepts in Action: A Practitioner's Toolkit*. Stanford, CA: Stanford University Press.

Winston, Jerome A. 1993. "Performance indicators: Do they perform?" *Evaluation News and Comment* 2 (3): 22–29.

——. 1999. "Understanding performance measurement: A response to Perrin." *American Journal of Evaluation* 20 (1): 95–100.