

**The Role of Government and the Government's Role in Evaluating Government:
Insider Information and Outsider Beliefs**

**Rod Dobell
University of Victoria**

**Panel on the Role of Government
Research Paper
August 12, 2003**

TABLE OF CONTENTS

Executive Summary	3
Introduction	9
Four themes in a changing world	
Backdrop: four decades of experience	
Definition and scope of review	
Changing context for evaluation – complexity, uncertainty and post-modernity	18
Systems approaches – the new sciences of complexity	
Post modernism	
Decline of deference; rise of rights; demands for involvement	
Systems response	
Central agency guidance – recommendations for practice remain positivist	24
OECD	
UK	
Canada	
Ontario	
Use of evaluation – enlightenment and influence are more important	32
Theory moves on – interactive deliberation, participatory integrated assessment	35
Academic literature on evaluation – US	
Participatory trends in European governance	
Boundary work and new criteria for effective assessment processes	
Importance of the ‘forum’	
Participatory evaluation within representative government	44
The concern	
Some illustrative possibilities	
Conclusions	49
Bibliography	54

Executive Summary

The conclusion of this paper is that the legitimacy and credibility of government can be better restored, and the social purposes pursued by government better achieved, through systems – including evaluation systems – that promote cooperation and trust rather than those that rest on competition and audit.

This conclusion flows from an attempt to tease out what might be learned from three or four decades of experience in the pursuit of systematic evaluation structures designed to support rational government decision-making based on compelling objective evidence. More particularly, the objective of the paper is to examine the last dozen years of work and literature on evaluation in an attempt to speculate on an appropriate role and orientation for evaluation activities in the government of Ontario in a nominal target year 2015, a dozen years in the future.

Over the past dozen years, a substantial shift in the context for thinking about government activities and governance more generally is evident:

- Human and natural systems are increasingly seen as complex, uncertain, highly interdependent and subject to continuing change, surprise and limited control;
- Epistemological convictions are shifting; postmodernism is increasingly influential; deference to allegedly scientific expertise is diminished; elite claims to objective evidence or ‘sound science’ are increasingly seen as tactics of power, not as grounded in superior knowledge; the extent to which ‘facts’ and knowledge are socially construed, conditional on a variety of unexamined (often unrecognized) conceptual choices is increasingly admitted;
- Citizen expectations of voice, participation, influence are sharpened, and local power to pursue realization of those expectations is increased, in substantial

part through networking opportunities opened up by new information and communications technologies.

- The problems of management arising from the perverse incentives and goal displacement associated with invalid or dysfunctional performance measurement systems are increasingly a matter of concern.

Faced with the emergence of these themes three or four decades ago, thinking in systems sciences and cybernetics moved toward possible responses in governance, envisaging some decoupling of governance structures, in some cases along lines described by the 'propose-dispose' model of Schon, advocating substantial attenuation in the control role of formal government and the rise of self-organizing community structures. And indeed structures and roles of government did evolve in response to these concerns, but more in the direction of decentralization to markets and deconcentration of central government. The image of 'steering, not rowing' was made popular. Decentralization through privatization and marketization was widely pursued, and the spirit of that program continues through deregulation and similar initiatives. Recognition of the challenges of interdependence also blossomed in the emphasis on 'horizontality' and 'joined-up government', with promotion of holistic corporate thinking. But in what remained of government roles not amenable to privatization, there was little let-up on concerns for control, and a massive increase in emphasis on accountability through increasingly elaborate systems for performance measurement and audit. Vigilant taxpayers envisaged minimal government presence policed by formal reporting and accountability frameworks entrenched in modernized government.

This emphasis remains the guiding star in most central agency guidance to public servants on the question of evaluation of government activities and reporting of

government performance, to the public and to legislatures. Accountability through increasingly comprehensive systems or through 'value for money audits' using prescribed comprehensive accounting methodologies is increasingly articulated, and departmental responsibilities for exhaustive review of their activities repeatedly emphasized as part of an overall obligation for reporting to ministers and legislatures.

It is, it might well be argued, a good thing for ongoing operational management itself to entrench evaluation activities in this manner (though it does also suggest relying on a classic politics/administration dichotomy to a greater extent than is warranted). Indeed we might argue that in Canada we have largely taken for granted the need to 'mainstream' such continuing cyclical evaluation activities, at least within operational management.

Experience suggests, however, that this aggressive agenda for evaluation in the 'modernizing government', New Public Management context is not succeeding in providing a foundation for executive politicians to deal with substantive change in significant policy orientations, any more than did the very closely related PPBS, MBO, MBR initiatives of forty years ago.

There are hints in contemporary literature on evaluation that thinking is returning to the alternative approach to dealing with complexity, uncertainty and the need for social decisions, namely through more thorough-going subsidiarity and devolution. Looking at the recent literature on evaluation theory, one sees recognition of post-modern ideas on the limits to knowledge and concerns for human flourishing or betterment as a goal of government programs reflected in suggestions for realist synthesis, responsive evaluation, empowerment evaluation, deliberative democratic evaluation, and many other variations. Looking to the literature on evaluation use,

one sees a broadening from concern with the use of evaluation findings as a flow of facts into decisions, to concern with the influence exerted by evaluation undertakings – that is, to the personal or organizational enlightenment or transformation that might flow from the learning associated with evaluation activities.

This thinking leads on to interest in participatory integrated assessment and ‘boundary work’ as sustained interactive participatory processes through which knowledge is co-produced and collective decisions arrived at through deliberative exercises. Rather than seeing a flow of objective scientific evidence as a basis for formal political judgments, one sees the emergence of collective intentions and public policies flowing from the interactive deliberations of participants dealing simultaneously with the negotiation and interpretation of the meaning of underlying information (understood through many ways of knowing) and the implications of collective intentions translated into individual action through compliance within those same participating communities (reflecting allegiance to the outcomes of legitimate deliberative processes).

As against the Hayekian response of decentralization to markets, coupled with continuing accountability through formal audit mechanisms, then, this approach reflects a Habermasian search for communicative rationality through subsidiarity and devolution to community-based groups, coupled with accountability through social roles. It reflects the suggestion that after three or four decades attempting to implement precise formal accountability mechanisms, we might try experimenting instead (or at least also) with the disciplining of individual discretion through internal means. That is to say, we might try the Friedrich side of the classic Friedrich-Finer debate, advocating education, suasion and cultural influence as the means to achieve

responsible administration in an uncertain world in which confident monitoring of individual compliance with collective intentions is impossible.

The conclusion is that in the year 2015, the role of evaluation in the Government of Ontario should be focused on:

- Effective knowledge management systems emphasizing tacit knowledge, not information and communications technology, in which MPPs might serve constituents as knowledge brokers;
- Assurance of accessibility (that is, truly effective access) for citizens to information in government hands, again facilitated through strong public service support for the efforts of MPPs in serving constituents directly, and perhaps through legislative task forces or roundtables;
- Open analytical and procedural support by the public service for interactive deliberative processes of shared decision-making in communities, particularly through imaging, visioning and simulation capacities, and participatory integrated assessment processes;
- Provision of fora – safe places for ongoing participatory involvement of citizens from the full range of interest groups or communities in shared governance.

This approach amounts to accepting the need for ongoing formative evaluation and performance monitoring as part of continuing routine management responsibilities supported by audit methods as usual, while concentrating evaluative activities on summative assessment in deliberative democratic fora for purposes of learning from the past in order to face the future differently. In the view of one British observer, the central conclusion is that we do not need more ‘modernizing government’, we need some ‘democratizing government’. The approach recommended here amounts to questioning whether, in a diverse society in a complex world, we need formal

accountability and policing of principal-agent relationships as much as we need increased trust in the legitimacy and reciprocity offered through enhanced opportunity for engagement in a more cohesive civil society.

The orientation for evaluation in the coming years should be “up the Arnstein ladder”, beyond a *pro forma* flow of government information outward to citizens, and toward more equal opportunity for citizens to be involved in truly participatory assessment of new orientations for policy formation and reporting of government performance in pursuing them. Recent Ontario initiatives in the direction of eDemocracy may offer openings for such a strategy. Initiating such developments in a manner that is compatible with Parliamentary democracy within a continuing federal structure is a serious challenge. But that is the challenge that must be faced, if an accepted role for government, and a credible approach to the evaluation of government roles, are to be found within the current mood and thinking of Canadian citizens.

Introduction

“Ideas in good currency, as I use the term here, are ideas powerful for the formation of public policy....By the time ideas come into good currency, they often no longer reflect the state of affairs...one of the principal criteria for effective learning systems is precisely the ability ... to reduce this lag so that ideas in good currency reflect present problems...it is by no means entirely inaccurate to say that government agencies are memorials to old problems”. (Schon, 1971, chapter 6).

Three or four decades of experience in the pursuit of methods for systematic evaluation in support of rational government decision-making based on compelling objective evidence should offer some lessons on which to draw in thinking about future development of analytical capacity and evaluation efforts in government. More particularly, the objective of this paper is to examine some of the literature from the last dozen years of work on evaluation in an attempt to anticipate what might be an appropriate role and orientation for such work in the Government of Ontario in a nominal target year another dozen years ahead, say 2015.

To anticipate the broad conclusion, this review suggests that we ought to move on from the current emphasis on accountability through formal performance measurement, performance management and performance reporting, and the current understanding of evidence-based decision as a flow of objective analytical results into a political decision-making process. Instead we should base concepts of evaluation as well as policy formation on sustained interactive deliberative processes in which citizens, representatives and officials are all engaged, on an ongoing adaptive basis. More specifically, the conclusion is that for well-established routine programs in which the underlying conceptual choices have long since been made, mental frames and belief systems are held in common across all those involved, and both agreement on objectives and the perceptions of surrounding environmental features can be taken as stable, the conventional canons of comprehensive audit and theory-based

evaluation centred on experimental or quasi-experimental methods can provide management on-going formative appraisal, and a framework for regular public reporting, but that this is not the case for the appraisal of significant programs or policy initiatives. On these, a body of ideas whose origins can be seen decades ago has been moving into good currency, reshaping notions of evaluation away from positivist analysis and toward participatory approaches recognizing many conceptual frames, many ways of knowing, and diverse perceptions of relevant consequences. In the last dozen years of academic literature, such ideas have been increasingly dominant. Now in the last few years we begin to see some central agency guidance to government officials moving in a similar direction, alongside the continued emphasis of treasury, finance or management documents on the standard efficiency-oriented audit-based approaches. The notions of assessment and appraisal centered in citizen engagement and participatory interaction have been moving strongly into good currency in the best official circles. It is the argument of this paper that this development should shape the perspective on evaluation in the Government of Ontario in the future; this probably means linking evaluation activities closely to emerging vehicles for E-democracy.

The underlying reasons spring from dramatic ongoing changes in the context in which governments function, and appraisal takes place.

Four themes in a changing world

There are at least four distinct and important underlying themes relating to performance reporting, audit and evaluation on which any panel undertaking a review of the role of government in the early 21st century should reflect¹. Having reflected on them, the challenge will be to articulate a stance that reconciles

¹ Dobell (1997) touches on some of relevant background.

recognition of these theoretical observations with the pragmatic and practical necessity for government to draw conclusions and take action in a timely and decisive manner, in the public interest, on behalf of citizens who cannot all be consulted individually, are often not informed, or interested, even though they should be concerned².

The first theme is that having to do with change and indeterminacy – the problem of attempting to evaluate programs or assess performance in the context of complex systems and profound uncertainty. When circumstances are changing significantly, continually, and unpredictably, it is not clear how to judge the effectiveness of good ongoing program designs that turn out badly from time to time. Or bad decisions that luck out through no virtue of their own. Evidence-based decisions are simply difficult in such settings. And audit or evaluation even more so, despite the apparent promise of automated management systems generating massive collections of records, but not yet much information.

The second, and closely related, theme is that which emphasizes the social construction – or at least socially construed character – of knowledge (or at least knowledge of social realities). Beliefs and perceptions vary widely, particularly when they address the distribution of risks and benefits associated with mysterious technologies or complex systems (Beck, 1992). So do values. And it is argued that the fact-value dichotomy, on which the politics-administration or steering-rowing dichotomies (or indeed the whole notion of responsible evidence-based decisions

² This inescapable necessity to act as agent for others who cannot be consulted raises many dilemmas, as noted briefly below. Among other things, it means that one cannot be content with following Putnam (1987) in resting moral justification simply on personal conviction and intuition as implied in his use of the Wittgenstein phrase ‘This is where my spade is turned’ – i.e., this is where I hit bedrock, where I can dig no deeper to examine the reasons or justifications for my beliefs. But that is a discussion for a different place. Some of the issues were sketched in an address to a group of senior executives in the federal public service some time ago (Dobell, 1989).

flowing from rigorous independent 'sound science') has collapsed (Putnam, 2002). How are differing perceptions, values and belief systems all to be brought together in systems that emphasize sound science in programs resting on evidence-based decisions? Further on in the metaphorical policy cycle, performance of policy rests on individual interpretations of the texts by a myriad of officials, and on the exercise of individual judgement in carrying out action. The agency of public servants, program clients and other citizens is involved. For this reason the question of access to performance data and program assessments becomes crucial, but also heavily contested, pitting competing interpretations and democratic rights to participation against personal rights to privacy and perhaps governments' rights to confidential advice.

The third theme is the growing public demand for involvement and participation in decisions on programs, in both design and management. And demands not just for voice, but also influence, or even veto. It is clear that those affected by government decisions have some right to be informed, perhaps consulted. Do they have the right to insist on consultation until they are persuaded? If program designs are not based on consensus among those affected, what is to be said in attempting to evaluate the outcomes? If the objectives have not been agreed, by whose criteria will (or should) programs be judged? Can it be argued that information and communications technologies will increasingly give citizens the means to judge directly, and choose accordingly, or will the vaunted openness of governments in the information age mean simply the emergence of a new elite, one that commands the resources to cut through the burden of data to the specific nuggets of information that advance their particular interests?

In all this, problems of dysfunctional measurement systems and goal displacement – the fourth theme – are crucial. The problems are not so much those of mis-measurement and mis-reporting on programs and program impacts (though these are significant) as they are in the impacts on incentive systems and behaviour for officials or clients or, in the increasingly common case of co-responsibility, both. (“What gets measured, gets done.”) This growing problem of using numbers because they are there (“in the absence of anything better”), measuring throughput or activity levels as if they were results, thus mistaking operational efficiency for program effectiveness, has been addressed in popular terms in a recent Massey Lectures series by Janice Stein, published as *The Cult of Efficiency* (Stein, 2001). The problem is pervasive; it needs more recognition generally, and certainly much more careful examination in the teaching of the new public management.

All four of these themes – complex systems and profound uncertainty; post-modernism in governance and public administration (French, 1992); participatory democracy; and the tendency to rely on dysfunctional performance measurement and assessment systems – raise theoretical challenges of great force in attempts to establish conceptually satisfying ongoing systems of evaluation, audit or performance measurement. Yet it is also necessary to come to timely conclusions, at least on a provisional basis, in order to take action to run a government; consultative processes can easily become a barrier to constructive solutions. (See Coglianesi (2002, 2003) for some expressions of concern on this point, as well as Coglianesi (2002b) for a salutary reminder that the efficacy of all the procedural reforms discussed here should themselves be the subject of thorough empirical analysis and evaluation.) The problem, from one perspective, is to find practical means to reconcile these conflicting theoretical considerations within general procedures, compatible with parliamentary democracy, that will be accepted as legitimate in particular cases.

Background: Four decades of experience

This paper takes a slightly longer view than usual of the evolution of thinking on questions of assessment of government activity (though not the full historical view sketched, for example, by Lindblom and Cohen (1979) or Poovey (1998) and other students of the philosophy of science). It includes brief reference to the production and release, on behalf of the Government of Canada, of the Guides to Operational Performance Measurement Systems, Program Evaluation, and Benefit-Cost Analysis, as well as other federal government manuals on systems for monitoring efficiency and effectiveness from the 1960s onwards³. It builds in particular on a recent review (Dobell, 1999) of Douglas Hartle's work in this field. That review argued that one should see an important underlying thread leading from an initial confidence in expertise and evaluation as technical craft to a later reliance on openness and public scrutiny of evaluation work, as part of a process in which public access to evaluation activities within government would support a more informed democracy.

Problems of interpretation are always present, however. The theme of an old paper – “If politics is theatre, then evaluation is (mostly) art” (Dobell and Zussman, 1981) – has since been developed to recognize more substantial problems of performance practice (see, in a different context, Taruskin, 1995) and post-modernism in attempting to assess whether program mandates have been faithfully and consistently realized, or whether excesses of compliance (excessive fidelity to the text, as assessed by auditors of authenticity) have led to failures in realizing the greater effectiveness that could be achieved through the capacity of discretion to respond to variety (Scott, 1998, esp Ch. 9). Problems of evaluation and accountability arising in the increasingly complex inter-organizational and inter-governmental (and hence cross-cultural) institutional

³ Muller-Clemm and Barnes (1997) offer a more detailed historical review.

structures through which programs are delivered were reviewed briefly by Dobell and Bernier (1999), and have been elaborated substantially in Kickert et al (1997) and Windrum and de Jong (2000) – though in sociology this literature goes back much further, at least to the early 1970s. But much more thought needs to be given to problems of interpretation and negotiation of meaning in societies characterized by deep diversity. Evaluation issues must be seen in the larger context of accountability concepts appropriate to a less scientific and positivistic, more diverse, subjective and constructivist, social world.

A review of practice across ten countries in undertaking evaluation work in this more general context – but specifically with respect to evaluation of social responses to global environmental risks – is contained in the chapter by Dobell, van Eijndhoven and Wynne in the report of the Social Learning Group (2001). That review surveyed four generations of evaluation theories, reflecting some recognition of problems of uncertainty and the need to respond to claims for participation. Written in the mid-90s, however, it simply took note of the ‘constructivist’ 4th generation approaches of Guba and Lincoln (1989) without articulating a response to the challenge posed. In some fashion, it may be seen as the task of the present paper to formulate some more usable response in dealing with the design of evaluation processes for governments.

Conclusions from the work of the Social Learning Group on knowledge into action, as summarized by Jäger et al (Social Learning Group, Chapter 21) led into more general work on assessment (Clark and Dickson, 1999) which continues now with the surge of interest in the flow of science into policy, the boundary work that goes into maintaining the distinct scientific and political cultures in the borderlands where these meet, and the boundary organizations developing to host such processes (Guston *et al*, 2000).

This review begins from this four-decade span of literature on program appraisal and evaluation in a government setting, and attempts to extend the basic concepts to reflect the realities of contemporary governance in Ontario while articulating principles consistent with the aspirations of increasingly active civil society organizations seeking to participate in a broad range of collective decisions in the face of profound uncertainty.

Definition and scope of review

Definitions of evaluation abound. The OECD Public Management Service has compiled one list (OECD, 1999) But it is worth emphasizing here that this paper is built on a significant shift of emphasis. Geva-May and Pal offer an example of the extreme away from which we want to move. Their definition of evaluation suggests that

“evaluation uses strict and objective social science research methods to assess, within various organizational structures, ongoing programs, personnel, budgets, operating procedures as well as to influence the adoption and the implementation of a policy or program” (Geva-May and Pal, 1999, p. 11).

Here we want to move away from reliance on ‘strict and objective social science research methods’ just as we want to move away from strict adherence to rigorous accounting standards or indeed away from any privileged authority attributed to ‘sound science’ as the preserve of any one mode of interpreting experience.

For present purposes, then, we can follow a more expansive practice and consider evaluation to be any study – or more generally process – designed and conducted to assist some audience to assess the merit, worth or shortcomings of some object or some coordinated set of activities directed at achieving expressed goals or purposes. (Here, if one thinks of a spectrum running from the most concrete prospective

(usually) project appraisal or project evaluation through retrospective (usually) evaluation of major programs to prospective (usually) evaluation of strategic policies, this paper is concerned with the latter portion of the spectrum.)

But it is a central feature of this discussion that in general we cannot view such evaluation activities as discrete, one-time events occurring in a separate stage of a policy cycle. Rather, in an adaptive management context, a learning context, we see evaluation absorbed within a seamless iterative process of intervention, appraisal and adaptation. At one level, this process may be entrenched within ongoing operational management dealing with what Sen calls the 'engineering' side of economics and resource allocation; at a more general level, it is undertaken within what Sen labels the 'ethical' dimension of economics and social decisions, where the full political challenge is addressed within participatory processes (Sen, 1987).

The consequence of this integration of evaluation within processes of collective decision-making is that a forward-looking discussion of the role of evaluation in government is inseparable from re-consideration of the role of social science research in policy formation, or indeed even more generally the nature of knowledge in the determination of collective action. This is unfortunate, but it seems consistent with the argument that there are no short-cuts, no resolution of basic social issues short of a full discursive process.

In other words, this paper starts from the premise that a process of evaluation is one aspect of the more general reliance upon expertise and expert advice in governance. This general problem has been addressed recently in a wide range of studies and papers, discussed below, dealing with the use of research or science by governments or in public policy more generally; this paper draws on that new and rapidly

expanding literature for lessons and guidance applicable specifically to the design of evaluation processes for governance in the future.

A Changing Context for Evaluation

“The old paradigm of scientific discovery (‘Mode 1’) characterized by the hegemony of disciplinary science, with its strong sense of an internal hierarchy between the disciplines and driven by the autonomy of scientists and their host institutions, the universities, [is] being superseded – although not replaced – by a new paradigm of knowledge production (‘Mode 2’) which [is] socially distributed, application-oriented, trans-disciplinary and subject to multiple accountabilities.” (Nowotny et al, 2003, p. 1)

“Present-day societies are characterized not only by pluralism and diversity, but also by volatility and transgressivity (in the sense of individuals, organizations and cultures acting beyond their traditional boundaries). The co-evolution of science and society has led to increasing complexity, unpredictability and irregularity in both spheres. Post-modern society has both a new perception of uncertainty and new means of dealing with risks.” (Salomon, 2001, p. 585)

Thus it is argued that the context for evaluation has changed dramatically over the last decade, at two levels, at least.

Systems approaches and dynamical systems: the new sciences of complexity

First, understandings and beliefs about the nature of the social and natural systems in which we live, and in which governments must formulate collective decisions, have changed dramatically. Uncertainty, complexity, limited capacity and limited controllability are all features that demand unprecedented attention in attempting to develop useful representations of the world around us. (See Waldorp (1992) or Casti (1994) for accessible survey descriptions of the ‘new sciences’, and Gunderson, Holling and Light (1995) for one extended attempt to link the evolving ideas in study of natural systems to corresponding concerns in complex social and economic systems,

or Berkes and Folke (2000) for another. Ongoing commentary can be found on the Resilience Alliance website at <http://www.resalliance.org/ev.php>.

Post-modernism

Second, the epistemological climate has seen substantial questioning of the 'enlightenment project'; the rise of post-modernism, or at least the recognition of the substantially context-dependent, socially constructed character of much understanding of the world (whether or not one wishes to deny that there is a world 'out there' at all) is the overarching challenge to be faced in reflecting on the flow of knowledge into decisions and action⁴.

Harvard philosopher Hilary Putnam has addressed the issue extensively, most recently in a series of lectures and essays, the most recent of which are collected under the title *The Collapse of the Fact-Value Dichotomy*.

"What we cannot say – because it makes no sense – is what the facts are *independent of all conceptual choices*." (Putnam, 1987, p.33; emphasis in original)

"Mundane reality looks different, in that we are forced to acknowledge that many of our familiar descriptions reflect our interests and choices." (Putnam, 1987. p.37)

"Thus, to come to think without these dogmas [the fact-value dichotomy and the analytic-synthetic dichotomy] is to enter upon a *genuine* post-modernism – to enter a whole new field of intellectual possibilities in every important area of culture [and economics]." (Putnam, 2002. p. 9; emphasis in original)

To a great extent, the conventional approach to evaluation, and more particularly computational procedures to support comprehensive auditing or value for money

⁴ The distinctions can be illustrated by reference to the art of the umpire, quoted by Schon and Rein (1994):

"I calls them as they is" (positivist)

"I calls them as I sees them" (post-positivist, or perspectivist)

"Until I calls them, they isn't" (post-modern, or constructivist).

examinations, rest on the possibilities denied in this post-modern framework, namely opportunities for an objective factual analysis built on 'sound science' and professional expertise, providing the evidence on which could be based the difficult political judgements necessarily reflecting subjective values. The alternatives outlined below propose more general lines of inquiry (in what Putnam describes as the Deweyian sense of the word).

"As John Dewey urged long ago, the objectivity that ethical claims require is...the ability to withstand the sort of criticism that arises in the problematic situations that we actually encounter..." (Putnam, 2002. p. 94)

"Inquiry in the widest sense, that is human dealings with problematic situations, involves incessant reconsideration of both means and ends;...if resolving our problem is difficult, then we may well want to reconsider both our 'factual' assumptions and our goals. In short, changing one's values is not only a legitimate way of solving a problem, but frequently the only way of solving a problem." (Putnam, 2002. pp. 97,98) (learning through experimentation and discussion)

"We do know something about how inquiry should be conducted...I mentioned the principle of fallibilism (do not regard the product of any inquiry as immune from criticism), the principle of experimentalism (...) and the principles that together make up what I called the 'democratization of inquiry' [including discourse ethics]....we need no better ground for treating 'value judgements' as capable of truth and falsity than the fact that we can and do treat them as capable of warranted assertibility and warranted deniability." (Putnam, 2002. 110)

In a remarkable address to senior officials in the Canadian federal government a decade ago, Richard French called attention to the way in which such post-Enlightenment developments were pushing government in Canada toward process concerns, toward a 'marked resurgence of expressive or symbolic politics' and 'the abandonment of analysis in favour of positioning' (French, 1992). This present paper may be seen in part as a delayed attempt to address the adjustment of strategies for evaluation in light of the developments identified by French.

Decline of deference, rise of rights

Third, beliefs about appropriate citizen roles, expectations of voice, rights to be heard have evolved sharply. So also has the capacity to give expression to those expectations, and indeed to give force to the demands.

From such a background emerge arguments for more participatory approaches to evaluation. But such arguments stem not only from public pressure; they also enjoy increasing academic support.

“The solution is neither to give up on the very possibility of rational discussion nor to seek an Archimedean point, an ‘absolute conception’ outside of all contexts and problematic situations, but ... to investigate and discuss and try things out cooperatively, democratically, and above all *fallibilistically*”. (Putnam, 2002. p. 45; emphasis in original)

A particular feature to note, however, of relevance to questions of governance in the future, and the search for equality in particular, is the fact that in a multicultural society of deep diversity, it is particularly dangerous to assume that there will be agreement on the conceptual context within which rational discussion of evaluative claims can be made. Without such agreement, the responsibility to take decisions on behalf of others who cannot be presumed to be offering informed consent poses many dilemmas as noted earlier (see footnote 2).

Systems response

In the face of all this, this collapse of confidence, the decline of deference, the claims for direct involvement, a couple of different responses are possible. One possibility proposed early on was the ‘propose-dispose’ design of Schon (1971), which envisaged an organic devolution of governance responsibilities to community-level self-organizing institutions operating within a general central framework. Another, of

course, is the development of a participatory democracy along the lines envisaged by former Prime Minister Trudeau in the early days of his administration. At about the same time, a notion much like the current image of adaptive management was introduced with Donald Campbell's idea of the Experimenting Society, in which social reforms were to be regarded and appraised as experiments (Campbell, 1969).

In fact the response generally took a somewhat different turn, a backlash against government, a conservative swing toward more limited government. Rather than the decentralization of government through devolution of responsibilities and development of a Habermasian deliberative democracy, the change in direction was towards a decentralization to markets, a Hayekian vision of smaller government replaced by transactions among individual economic agents rather than interaction among empowered citizens⁵. Along with this came growing political support for the New Public Management, with a focus on formal accountability mechanisms and the challenge of principal-agent problems.

Thus, in this context the roles of government and the public service have changed, the instruments of governance have changed, and the structure of government has changed. Delivery of services to citizens as clients has become an overarching concern. One can note, for example, the Ontario government's emphasis on "revolutionizing services to clients through single windows and computerized contacts" (Ontario, 2000). Even though the focus of government activity at the provincial scale is on the provision of services in health, education, social services, however, there is still much talk of 'steering, not rowing', of less emphasis on direct service delivery and more on contracting, partnerships, networking of voluntary

⁵ Interestingly, Culpepper (2003) suggests that the two lines of development (in the form of the somewhat similar 'Empowered Participatory Governance' and 'Market-Preserving Federalism') share both some important features and some crucial limitations.

sector organizations, and other inter-organizational coordination activities with civil society organizations. 'Distributed governance' has become a catch phrase; subsidiarity a touchstone. Concerns for formal accountability and control mechanisms mount as a result. Regulatory negotiation processes, results-based standard setting and alternative dispute resolution mechanisms replace some traditional regulatory responsibilities. More importantly, responsibilities for managing social risks become priority concerns, but without any corresponding institutional capacity in government to address them. The 'evaluands' in evaluation activities become ongoing partnerships or networks, ongoing decision processes, perhaps institutions or constitutions of nested institutions, more than discrete programs, thus posing major methodological challenges for evaluation and accountability.

Despite all these changes just sketched, there remains an ongoing body of straightforward performance measurement, management and reporting activity which represents the vast bulk of what is recommended in standard literature on evaluation. There is no point in this paper going over either the procedures or the institutional arrangements for all this work yet again; the relevant features have been summarized recently in a project for the Canadian Evaluation Society (Zorzi et al, 2002), and recent documentation has been surveyed by Segsworth (2003) for this panel. The following section simply takes note of some elements of continuing central agency guidance on this more instrumental and technical end of the evaluation spectrum.

In general, this guidance from treasuries and management boards is unambiguous about the approach in practice, which remains more technical than deliberative, and, where it ventures outside the bounds of inside science, is more consultative than participatory. But, as noted below, a different line of argument is beginning to emerge more visibly from other central agencies.

Central Agency Guidance.

“Officials and ministers must have reliable, relevant and objective information available if real improvements in the decision-making process are to be achieved.” (Government of Canada, Treasury Board Secretariat, Study of the Evaluation Function in the Federal Government, March, 2000, p.1)

“Thus the accuracy of transcription, which proclaimed the rectitude of the books, stood in for what could not be verified, the accuracy of the initial record of goods and transactions” [that is, the accuracy of observations] (Poovey, 1998. p. 56)

Christopher Pollitt has a succinct description of the wave of publication of European government guides to evaluation in the mid-1990s, following upon the North American reforms a little earlier, but a little prior to the most recent round of such publications in Canada. Writing in 1997, he comments on the surge of enthusiasm for evaluation among European governments, noting that

“The official guides make evaluation sound as though it were an essential basis for rational policy-making and programme management...All the guides mention the need to tease out and pin down the logic of a programme. They envisage the evaluator being involved in clarifying final objectives, intermediate objectives and operational targets, and in trying to establish which programme activities contribute to which programme objectives. All cite economy, efficiency and effectiveness as important evaluative criteria. ... All spend some time discussing different kinds of evaluation design, and the appropriateness of particular methods and techniques in particular circumstances. Most of them state or imply that the experimental method, with control groups, is the ‘gold standard’, though they also point out that it cannot always be applied.” (Pollitt, 1998)

Since then there have been a number of further developments, leading some governments much more toward concerns for openness and inclusiveness, and for participatory processes in general. Reference to these, along with the conventional guidance, will be sketched only very briefly here, beginning with the cross-national comparative work at the OECD.

OECD

The general thrust of thinking in many OECD countries is well reflected in a couple of recent publications from the OECD, embracing generally the standard 'scientific' approach to evidence-based decision. A background paper on improving evaluation practices (OECD, 1999) sets out a concise summary of key principles underlying a statement of *Best Practice Guidelines for Evaluation* published slightly earlier as a Policy Brief (OECD 1998).

More general concerns with accountability and control, including some anxiety about evaluability in the broadening public sector, are addressed in an extensive study of *Distributed Public Governance* (OECD, 2002) which is described as "concerned with the protection of the public interest in the increasingly wide variety of government organizational forms'.

At the same time, however, an extensive report entitled *Citizens as Partners: Information, Consultation and Public Participation in Policy-Making* provided the basis for another Policy Brief and a Handbook also entitled *Citizens as Partners* (OECD 2001)⁶. These documents set out the case for more extensive participation in processes of evaluation as well as policy formation. They also introduce a simplified version of the 'Arnstein ladder' (Arnstein, 1969), noting that varying degrees of citizen involvement and engagement are possible.

⁶ On reflection, this may seem a somewhat odd metaphor, implying a view of governments as autonomous entities with whom sovereign citizens must partner on some basis of mutual benefit, rather than seeing governments as one of many structures created by communities to coordinate their various efforts and decisions in some common interest.

The Green Book--UK

A good summary of contemporary central agency thinking on questions of appraisal and evaluation can be found in the guidance offered by HM Treasury in the United Kingdom, the famous 'Green Book', *Appraisal and Evaluation in Central Government* (HM Treasury, 1997). Interestingly, the current (undated) edition in use in 2003 has elevated the informal usage: the title now appears as *The Green Book: Appraisal and Evaluation in Central Government* (HM Treasury, 2003).

It is helpful to note the way in which the purpose of the work is characterized in this current edition. The Preface begins

"The Government is committed to continuing improvement in the delivery of public services. A major part of this is ensuring that public funds are spent on activities that produce the greatest benefits to society, and that they are spent in the most efficient way.

The Treasury has, for many years, provided guidance to other public sector bodies on how proposals should be appraised, before significant funds are committed – and how past and present activities should be evaluated. This new edition incorporates revised guidance, to encourage a more thorough, long-term and analytically robust approach to appraisal and evaluation. It is relevant to all appraisals and evaluations. Appraisal, done properly, is not rocket science, but it is crucially important and needs to be carried out carefully. Decisions taken at the appraisal stage affect the whole lifecycle of new policies, programmes and projects. Similarly, the proper evaluation of previous initiatives is essential in avoiding past mistakes and to enable us to learn from experience. *The Green Book therefore constitutes binding guidance for departments and executive agencies.*" (HM Treasury, 2003; emphasis added.)

It is interesting also to note the key points that the last two editions have flagged as revisions. Commenting on changes from the 1991 edition to 1997, the Foreword to the 1997 edition says

"This edition...provides more material on evaluation. It gives greater emphasis to the appraisal and evaluation of environmental impacts. It takes account of developments in the treatment of these impacts and other costs and benefits which are not easy to value. It takes account of developments in the use of private finance and, related to this, provides a more thorough coverage of the treatment of risk and uncertainty. The

section on industrial and regional programmes has been extended to cover a broader range of programmes aimed at raising economic activity.

The guide confirms the use of a 6% real public sector discount rate in most circumstances.” (HM Treasury, 1997, p. vii)

From 1997 to the present edition, the highlighted revisions are described in the following way.

“First, there is a stronger emphasis on the identification, management and realization of benefits – in short, focusing on the end in sight, right from the beginning. Secondly, the new edition “unbundles” the discount rate, introducing a rate of 3.5% in real terms, based on social time preference, while taking account of the other factors which were in practice often implicitly bundled up in the old 6% figure. *In particular, the new Green Book includes, for the first time, an explicit adjustment procedure to redress the systematic optimism (“optimism bias”) that historically has afflicted the appraisal process.* Finally, there is greater emphasis on assessing the differential impacts of proposals on the various groups in our society, where these are likely to be significant.” (HM Treasury, 2003, p. v, emphasis added)

Thus fundamental social change is selectively mirrored in these shifts in the orientation and emphasis to be captured in the computation processes underlying the flow of information to senior officials and ministers in support of decision-making. What might in some settings be considered rather basic social and political choices are here simply embedded in the technical guidance.

Outside of HM Treasury itself, however, there is in the UK Central Government a substantial, and somewhat different, flow of guidance dealing with questions of consultation and the appropriate roles for science advice. See, for example, *Policy, Risk and Science* (United Kingdom, 2000); *Guidelines 2000: Scientific Advice and Policy-Making* (United Kingdom 2001); *Consultation Guidelines* (for written consultation) and *In the Service of Democracy* (for e-consultation), the first two of which deal with science advice as expertise, but the last two of which emphasize the growing importance of sustained two-way interaction with a broader set of participants.

Canada

The story in Canada is a little different; the Canadian government has perhaps backed away somewhat from the ambition to offer 'binding guidance' to departments. There is no need to repeat here the voluminous literature that describes the evolution to the current focus on results-oriented management and accountability. One can note the extensive materials on the Evaluation website of the Treasury Board Secretariat (at www.tbs-sct.gc.ca/eval), setting out the new policy on evaluation, effective April 1, 2001, as well as an account of the work of the new Centre of Excellence on Evaluation, created at the same time. All of this is structured within the government's overall management philosophy as set out in the document *Results for Canadians*. Within this general political statement is constructed the Results-based Management and Accountability Framework (RMAF) through which deputies and managers are required to pursue the new evaluation policy. Among the items on the website are the 1981 versions of all this, the *Guide on the Program Evaluation Function* and the *Principles for Evaluation*. Not on the website, but relevant to any assessment of the evolution of approaches to evaluation are the earlier documents on *Operational Performance Measurement Systems* (Canada, 1973) and *Principles of Performance Measurement* (Canada, 1978) all of which represent attempts to implement the underlying objectives of the Reform of the Estimates initiatives leading to the adoption of the Planning, Programming Budgeting System in the late 1960's. Dobell (1999) attempts to make the case that there is little progress evident in the move from this early work on management by objectives/performance measurement/management by results as components within the overall ambitions of the PPBS work, and the current initiatives recast in the language of comprehensive audit. There is no need to pursue that argument here. Segsworth (2003) provides a more extensive summary.

A more recent commentary (Canada, 2003) emerging from a meeting of the community of heads of evaluation in Ottawa emphasizes two features worth noting. The first is the repeated reference to an existing or anticipated shift to a 'culture of reallocation' as setting the crucial context for evaluation work in the near-term future. The second is the explicit, but unexamined, premise that the audience for evaluation results is the executive committee of the department – that departmental analytical work in evaluation provides information to support decision-making inside those departments⁷. The case being made in the present paper is, of course, rather different: it is that means must be found to ensure that this analytical capacity can provide a foundation for inclusive processes of deliberation aimed at reconciling the diverse beliefs of citizens outside government with respect to the merits of government programs.)

At the same time, the development of ideas around citizen-focussed government has led to a number of guidelines emphasizing consultation and participation in the formulation or appraisal of government policies. These can be found in a variety of places on the website of the Privy Council Office, where one approach has been enunciated by the former Clerk, who said "Better consultation processes by government will result in better policies that are better understood and better meet the priorities of Canadians" (Cappe, 2000) .

More ambitiously, in a 1998 address, the previous Clerk to that articulated the goal as follows, under the heading "**I am talking of citizen engagement**" (her emphasis):

"Citizens wish to relate to their democratic and public sector institutions in new and different ways. They are no longer satisfied to participate in an election every four or

⁷ The theme of the meeting was 'Serving the Decision-making needs of the Executive Committee'.

five years. They want to have a say in the policies that affect them most. They want to be partners in shaping Canada's future.

Over the years, we have gained experience in using different ways of involving citizens -- from the provision of information, to the reporting of results, to major consultation processes. As we speak, there are 300 public consultations exercises under way across the Public Service of Canada.

We must now go beyond this to a new frontier and learn about citizen engagement. It is a two-way learning process between citizens and their democratic and public sector institutions. It involves trade-offs and a search for common ground. It is not easy; it is time consuming; it can be costly. But used appropriately and selectively, the results are worth it." (Bourgogne, 1998).

Ontario

The Ontario Government's Management Board has recently developed guidance to departments on performance evaluation, but has done so in the form of a Management Board Directive, which represents (or at least can be construed as) advice to Cabinet, and therefore not a public document, and therefore not accessible on the Government of Ontario website. Nevertheless, with the cooperation of the Management Board Secretariat (Program Management and Estimates Division) an examination copy of the March 2003 edition of the Guide to Program Evaluation (sub-titled Work-in-Progress) has been made available for review for purposes of this project. The Guide offers an excellent manual to support ongoing work on program evaluation in circumstances where activities are sufficiently stable and well enough defined to be amenable to the conventional instrumental approach. It does not attempt to address the more general public processes discussed below.

More general developments are sketched in the April 2000 report, *Transforming Public Service for the 21st Century* (Ontario, 2000). Of particular note in the present context is the reference to the experience of the Ministry of Natural Resources with its Lands for Life initiative, a planning process established to recommend options for the future of

the Crown land base of Ontario. In this process, three regional Roundtables were established, from which a number of recommendations were consolidated in a proposed land use strategy released as *Ontario's Living Legacy* in March 1999.

What is of particular relevance for the discussion below is the suggestion that this process established 'the basis of a new relationship' with groups directly concerned.

"Because this process was so open, it put huge demands on our (MNR) information base. We had to make sure that the information was in a format people could use. In the process we had to develop a whole set of new analytical tools..." "When the public knows as much as the public servants do, the relationship changes. We can never go back to the old mode where we give the public a peek at the information we have. It's now all out in the open." (Ontario, 2000, pp 32, 33).

While most of those in the NGO community would consider this characterization more than a little optimistic, the idea is central.

And it appears to have staying power. In preparation for the April 30, 2003 Speech from the Throne, the Government of Ontario launched a consultation process using Internet facilities. About 2500 people responded, although obviously the use in the development of the Speech from the Throne was more anecdotal than substantive. An April 30th, 2003 press release on the Premier's web site announced further steps toward the 21st century vision, including the following:

"Consistent with the belief that government exists to serve people, and not the other way around, the government will expand its use of the Internet to help bring citizens closer to their government. The goal is to ensure citizens' access to a wide range of tools and information that will enable them to participate more fully in the democratic process."

In summary, the point of this brief sketch of central agency guidance is simply to note that the conventional government prescriptions on evaluation practice have

thoroughly embraced and endorsed the positivist orientation grounded in what Poovey calls the 'figures of arithmetic'. Despite all the evident problems with dysfunctional performance measurement systems, perverse incentives, goal displacement (with consequent legislative and public inattention) and implementation problems experienced over the last thirty or forty years, the attempts continue to create a seamless flow of performance reporting that will conform to generally accepted accounting principles, provide rigorous and objective evidence for decision purposes, and make auditors-general happy.

But the problem is that experience over those same thirty or forty years suggests that all these findings and all this reporting seems not to help much with significant reappraisals or policy reorientations. The history of evaluation work is often characterized as one of 'unfulfilled promises'. Evaluation findings do not seem to shape big decisions about terminating or initiating major programs. The next section probes this question a little more deeply.

Use of Evaluation

"Policy makers rarely base new policies directly on evaluation results." (Weiss, 1999, p. 468).

"A conclusion that may be drawn from our examination of the extent to which evaluations led to learning is that only in exceptional circumstances do evaluations appear to play a part in significant changes in orientation of policy" (Furubo, 2003).

This concern has been a theme also in Canadian evaluation circles for well over three decades. The preferred response (in the early 1970s I used it extensively myself with analytical staff discouraged by the apparent failure of their briefing notes to carry the day with Ministers, Treasury Board or Cabinet) is to find comfort in the famous lines of Keynes about the slow spread of new ideas. Now that same comfort is given more professional standing by more explicit labels. Weiss (1998) speaks of the

'enlightenment' use of evaluation studies – the consciousness raising and improved awareness of issues flowing from the introduction of new ideas or professed findings into political discussion – distinguishing the instrumental use of evaluation findings in specific decisions from the general 'enlightenment' surrounding social science research work and evaluation studies. (See also Weiss (1999), Julnes and Holzer (2001), Perrin (2002) and Cummings (2002).) Patton (1998, p.225) speaks of 'process use', the impact of evaluation studies on the people and organizations that undertake them. Again going back to arguments familiar for decades in the futures studies or strategic planning literature, it is argued that it is not the explicit result that matters so much as it is the learning or tacit knowledge gained from the experience of going through the thinking process. Interestingly, Patton speaks of the inter-cultural character of such experience, and the benefits of the inter-cultural encounter. Though he leaves the impression that the learning is largely by those in the political culture, learning from those in the research culture, his observations lead on naturally to the broader body of writing now emerging on what is called 'boundary work', to which we will return later in this paper.

Kirkhart (2000) proposes an integrated theory of influence that includes both the potential use of evaluation findings and the potential impacts arising from the activity itself. She suggests that "this integrated theory of influence helps us to recognize that evaluation practice has had a more pervasive impact than heretofore perceived." (Kirkhart, 2000, p. 20)

Nonetheless, the additional dose of optimism associated with these perspectives, it must be emphasized, has more to do with the learning processes possibly associated with the activity (see also Preskill and Torres, 2000) than with any increased expectation of direct use of the numerical, factual or specific evaluation findings.

Parsons (2003) has a different way to express a similar notion, namely to suggest that the challenge for governments is not really modernization, it is democratization.

Furubo (2003) draws the same conclusion in more technical language, suggesting that “The information acquired from evaluations does not seem to be a major explanation for significant policy changes, but they are used in fine tuning and implementation in operative decision-making, where there is continuing increase in the evaluative information reaching policy makers, and it is being used to convey information related to what has been accomplished by agencies...with an orientation toward output and performance more than on effects or preconditions (context) for effects. This leads one to visualize a dual future development...”

“On one hand we have perhaps a tendency towards a more on-going, continuous stream of evaluative information, which perhaps will be channelled more or less directly to the administrative decision-makers from different systems. This kind of information is generated more or less in direct relation to the implementation of different activities. It gives the agencies, or more generally speaking the organizations responsible for the implementation, a key role in the production of such information [directed to formative evaluation]. On the other hand we have the development of other forms of evaluations, which have to be done on an ad hoc basis, and which more strongly than ever before today need to interact with a more general science production.” (Furubo, 2003, p.11)

Perhaps in summary it is worthwhile going back to the opening line of this section, drawn from the abstract of a recent article, to quote that abstract at greater length. It reads

“Evaluation has much to offer policy makers, but policy makers rarely base new policies directly on evaluation results. Partly this is because of the competing pressures of interests, ideologies, other information and institutional constraints. Partly it is because many policies take shape over time through the actions of many officials in many offices, each of which does its job without conscious reflection. Despite the seeming neglect of evaluation, scholars in many countries have found that evaluation has real consequences: it challenges old ideas, provides new perspectives and helps to re-order the policy agenda. This kind of ‘enlightenment’ is difficult to see, and it works best when it receives support from policy champions. Many channels bring evaluation results to the attention of policy makers, and they listen not only because they want direction but also to justify policies, to show their knowledge and modernity, and as a counterweight to other information. Openness of the political

system and a thriving evaluation community tend to make some nations more attuned to evaluation than others." (Weiss, 1999, p. 468)

More strongly, the balance of this paper goes on to develop the last line of the comment by Furubo just quoted above:

"On the other hand we have the development of other forms of evaluations, which have to be done on an ad hoc basis, and which more strongly than ever before today need to interact with a more general science production."

This reference to a more general science production is the link into the notions of co-production of knowledge, and communicative action, as basic to methods for social decisions through democratic evaluation or interactive democratic processes, as outlined in the next section.

Theory moves on

"...the plea from our end of the academic community would be to quit talking about plans and intentions and procedures, and get on with the work. And let the rest of the world in to see that work [openness and accessibility], so as to draw its own conclusions about how valid that work is [metaevaluations] and how effective are the programs appraised. Truth has never been effectively pursued or persuasively spoken in any other manner." (Dobell and Zussman, 1981, p. 427).

Over the past decade, much academic work on evaluation theory has moved substantially away from the positivist, program theory based models for evaluation described above, toward attempting to meet the needs of evaluation in a complex, uncertain world in which 'Mode 2' science (Nowotny et al, 2003) must be brought into collective assessments through open participatory processes.

Somewhat earlier, in 1990, Doug Hartle argued for an institutional reform that would pursue this goal through assignment of peak evaluation responsibilities to the Senate Committee on National Finances (though he later lost confidence in even this attempt to identify within existing government structures an appropriate forum for open,

accessible appraisal). More generally, his move from an internal technical orientation based on rigorous social science in evaluation to an outward discursive posture resting on faith in democratic process – an evolution that seems to me fundamental to understanding the development and future direction of evaluation in Canada – is traced in Dobell (1999).

New directions in evaluation: US

Much the same evolution can be seen in many accounts or surveys of evaluation theory in recent years. Stufflebeam (2000) offers an encyclopaedic survey of evaluation approaches, noting the substantial evolution of what he calls “social agenda/advocacy approaches” – from utilization-focused evaluation (Patton, 1978), to responsive (client-centered) evaluation (Stake, 1975; Wadsworth, 2001), constructivist evaluation (Guba and Lincoln, 1989), realistic evaluation (Pawson and Tilley, 1997), empowerment evaluation (Fetterman, 2001), and deliberative democratic evaluation (Floc’hlay and Plottu (1998); House and Howe (2000); Ryan and DeStefano (2000)). Inevitably, of course, there are also sceptical reservations about the possibilities for such empowerment and democratization, often based on concerns about the uses of power and the barriers to power sharing, along lines usually attributed to Foucault (see, for example, Gregory, 2000).

Participatory trends in European governance

Independently of this flow of (generally) American and British academic work, it seems, there is another important stream of European literature building up around European governance projects (see the White Paper on Governance (European Commission, 2001) as well as an associated working group report (Commission of the European Communities, 2001) on ‘democratizing expertise’ . An important survey of proposed frameworks and processes for the engagement of civil society in policy

formation and evaluation activities is given in the interim report *The Role of Civil Society* (Banthien et al, 2003), relating to the proposed European Research Area.

Because of their reliance on long-term developments in a context of complex institutional structures and unknowable natural dynamics, the Research and Technical Development (RTD) programs are seen to demand particular attention for evaluation purposes. Specifically with respect to evaluation, other surveys (for example, Windrum and de Jong (2000) or de Jong et al (2001), dealing with the evaluation of RTD programs through the project on Simulating Self-Organizing Innovation Networks (SEIN project)), review the broad evolution of evaluation theory with an eye to its necessary further development in the context of the complex dynamics of innovation networks and distributed governance.

This changing structure of government activities and consequent change in what is evaluated emerges within the changes in the expectations and beliefs against which government activities are evaluated, as noted above. Citizens, often marginalized groups needing the support of effective social programs, are identified as important 'users' of evaluation work, thus bringing a focus on 'empowerment' and 'betterment' directly into the design of evaluation criteria. Social beliefs and expectations increasingly recognize the complexity and profound uncertainty surrounding natural and social processes and collective decisions intended to influence interventions in them. There is growing respect for expectations of involvement, engagement, voice and influence on the part of individuals affected by government decisions: the legitimacy of decisions reached and activities undertaken without appropriate inclusive participation is increasingly questioned. There is growing acceptance also of the extent to which social realities are socially construed: even if there is a social reality 'out there', it is perceived very differently by different people in a society

characterized by deep diversity (Douglas, 1999), and the negotiation of consensus around what is observed and the meanings to be derived will be protracted at best. The mantra of evidence-based decision and reliance on 'sound science' is challenged by post-positivist perspectivist or post-modern constructivist belief-based systems.

Boundary work and new criteria for effective assessment

The criteria for assessment must change as a result. Evaluation must explore the resilience and adaptive capacity of activities assessed as ongoing processes; must consider the extent to which these processes realize goals of inclusiveness, transparency, robustness against shocks, achieve goals of continuous learning and process improvement, and constant innovation. More general criteria are discussed in the context of attempts to appraise the impact of analysis in support of policy formation by Clark and Majone (1993), Cash et al (1997), Clark et al (2002) and other work in the Harvard Global Environmental Assessment Project (Clark and Dickson, 1999). Beyond this, the whole process by which analytical work might support collective decisions, either in prospective appraisal for policy formation, or retrospective evaluation for ongoing learning or summative appraisal of merit, has become a field of study of its own, leading to much exploration of 'boundary work', 'boundary organizations' and negotiated hybrid outcomes, or 'boundary objects'.

These questions of boundary work, boundary objects and boundary organizations are only beginning to be explored in the social studies of science literature, and also in the emerging literature on global environmental assessment literature; as just noted, some of the current work comes out of Harvard as a follow-up to concerns about the flow of knowledge into action that resulted from the work of the Social Learning Group (2001). The work of Leigh Star and others from about 1989 onwards on 'boundary objects' as devices designed to help communication across the boundaries of

communities of practice forming into communities of interest, and particularly across boundaries from scientific communities to policy communities (Star and Griesemer, 1989); or the related work of Gieryn (1983) and others from around the same time on the 'boundary work' that particularly serves to demarcate the boundaries of scientific communities from the other allegedly less legitimate sources of advice into policy communities and to underscore the greater authority of credentialed expertise or 'scientific evidence'; or finally the more recent work on the 'boundary organizations' that might smooth, not just the flow of science into policy, but the co-production of evidence base, relevant knowledge and policy inferences through continuing interactive deliberative processes involving both scientific and other communities (see Eden, 1996); all explores the question of effective participatory processes bridging distinct communities or cultures.

St. Clair (2002) argues for the extension of the ideas emerging from this Harvard-centered work on policy to address global atmospheric risks, and the European work more generally on sustainable development policy, to address the social challenge of poverty reduction as boundary work. Again in her work, the legitimacy criterion calls for a thorough-going deliberative democracy. Nagy (2002) in her work with the OECD on the Aarhus Convention, offers principles and guidelines to involve 'concerned publics' and a general public to varying degrees, with claims of the former for participation in decisions that affect or interest them, claims of the latter perhaps limited to access to information.

In the somewhat distinct field of human-computer interaction, there is also a growing literature on this subject of cross-cultural work supported by boundary objects, some of it focussed on the anthropologist's concept of 'trading zones'. Some of that literature pursues the analogy with the task of translation, which leads into questions

of performance practice (excessive fidelity to the text? who is to be moved toward whom?) which leads in turn into the whole topic of the implications of post-modernism for public administration and public policy, particularly as processes of implementation attempt to interpret in the region the texts written at the centre. (See, for example, Chrisman (1999); Muller (1999); Arias and Fischer (2000); Thagard (2001), and again the insightful work of James Scott (1998).)

In yet a different setting, addressing primarily issues of social risk (Beck, 1992), work on assessment in the context of uncertainty and complexity has led to a flow of guidance on deliberative interactive processes (see the foundational document edited by Stern and Fineberg (1996) for the National Research Council in the US), and a review of the criteria for assessment of policy advice in boundary work in multi-tier institutions spanning multiple scales, as noted above. From all this current work on global environmental assessment processes and the role of boundary organizations comes the articulation of three core criteria in determining the success of assessment processes. These are salience, credibility and legitimacy (Clark et al, 2002). (And within the criterion of legitimacy, the European survey mentioned above identifies openness and transparency, inclusiveness, and interaction as determining features.)

With criteria of legitimacy, credibility, and saliency as critical features of the flow of evidence into policy formation, the process, not surprisingly, proves crucial. "Our studies have suggested that much about what makes some assessments more effective than others seems to be tied up with the *process* by which they were developed, rather than just the *product* itself." (Clark and Dickson, 1999). These criteria must be kept

centrally in mind in considering the effectiveness of proposed institutional arrangements and processes for evaluation of government performance⁸.

It is interesting to reflect on these goals in conjunction with what have been set out as three fundamentals of deliberative democratic evaluation (House and Howe, 1998): inclusion, dialogue and deliberation, and with concerns about an overly technical emphasis on in traditional assessment of government performance. Similarly, materials for a conference on Policy Learning and Distributed Governance hosted by the Canadian High Commissioner to the UK (the former Clerk of the Privy Council, Mel Cappe) in London June 5, 2003 emphasize ‘innovation, transparency and accountability’ as “accepted goods in government, forming a policy prism through which programmes, old and new, are evaluated”.

And indeed the terms of reference for the present panel call for consideration of the impacts of existing or proposed institutional arrangements for evaluation on “democratic values such as accountability, transparency and public participation”.

In many ways this is the same basic idea emerging in different words in different contexts. Weiss (1999) argues for sustained interactivity while extending the notion of ‘user’ to clients and civil society. Stern and Fineberg (1996, cited above) – in what

⁸ Interestingly, the concern with legitimacy as a central criterion for an effective evaluation process does not conflict – as it might seem initially to do – with the strongly argued posture of Kaplow and Shavell (2002), making the case that in legal policy-making, or more generally in public policy choices, the criterion of fairness or legitimacy should have no weight independently of its role in determining welfare or wellbeing. That case is well made; but here the concern with legitimacy is in the processes that shape and frame problem formulation itself, in the process by which voices are heard in bringing perspectives and ways of knowing to bear in understanding meaning and consequences in framing the discussion and the decisions faced, not in the criterion for choice among possible responses. That choice may appropriately be resolved on the basis of welfare criteria only; but the legitimacy of the process which led up to the framing of the choice is itself of independent, indeed crucial, importance.

must be viewed as a comprehensive and authoritative study – argue the need for democratic interactive processes. A wide range of recent contributions in the academic literature on evaluation argue the case for empowerment evaluation, participatory evaluation, democratic evaluation, as noted above. The central feature in all this is that interested citizens, clients and stakeholders are involved from the beginning of the process of agenda formation and problem identification, at the stage where what Putnam (1987) calls ‘concept selection’ occurs to determine the perspectives from which issues (and rules of evidence) will be defined and selected, through to the realization of policy intentions in individual action⁹.

The importance of the forum

The point here is simply to emphasize the way in which the conduct of evaluation is pressed away from the project appraisal end of the spectrum, where ‘facts’ are less contested and more easily negotiated, and technical methods work in the assembly of evidence, and toward the policy evaluation or institutional assessment end of the spectrum where reconciliation of conflicting frames or belief systems and negotiation of agreement on a common understanding of outcomes have nothing to do with accounting or computation and everything to do with communication, persuasion, influence and power – and hence everything to do with the setting for discourse.

All of this leads on to the argument so well set out in the summary volume by Schon and Rein, *Frame Reflection* (Schon and Rein, 1994). There the argument is for design rationality, with the key step being the possibility of achieving a communicative setting (the Habermasian vision of the ideal speech situation) in which participants

⁹ The *Study of Governance of the European Research Area* (2003) cited above provides an extensive discussion of alternative frameworks and processes. Toth (2001) undertakes a searching appraisal of the potential for participatory integrated assessment in that same context, with respect specifically to potential applications in the work of the European Environment Agency.

can reflect on the mental frames they bring to policy controversies and, more importantly, can come to appreciate the frames defining the perspectives of other participants in such controversies. This process of mutual reflection may be seen as an ethical obligation, or simply a pragmatic response to the need to resolve controversies stemming from the conflicting action frames shaping the views of the various parties. But in any case the process goes beyond a simple transfer of numbers and analysis, and demands a continuing forum or space in which mutually respectful discourse can continue.

Exploring in comparative context how such bodies build credibility and legitimacy is a major contemporary challenge. The question as to how community mapping, visioning or simulation capabilities might function as boundary objects to increase the effectiveness of the interactions of community representatives, scientists and technical advisors, and politicians or officials is both a central theoretical question and a practical issue of building capacity for more effective involvement of community participants in consultative processes.

Thus all these proposed approaches to assessment and decision touch on two crucial implications: the need for the public service to offer extensive analytical and technical support to the deliberative process, and the importance of the 'forum' as a safe and hospitable place, an ideal speech situation in which communicative rationality can be sought. This need for a 'forum' is felt at many levels, from the community-based boundary organization such as the spontaneously emerging marine resource centres in small coastal communities to the large-scale citizens forum or consensus conferences mounted by formal mainstream organizations such as the Public Policy Forum in Canada or the Technology Boards in the Netherlands or Denmark. This

emphasis on the forum itself moves beyond procedural rationality to the pursuit of communicative rationality as such.

An evocative expression of this interest in developing appropriate fora for the sustained involvement of citizens in participatory deliberative processes is the concept of the *agora* advanced by Nowotny *et al* (2001). They introduce the image of the *agora* notion as

“the problem-generating and problem-solving environments in which the contextualization of knowledge production takes place. It is populated not only by arrays of competing ‘experts’ and the organizations and institutions through which knowledge is generated and traded but also various jostling ‘publics’in its own right a domain of primary knowledge production.”.

Still more generally one can – perhaps must – address the question of the link between analytical style or evaluation approaches and the nature of democracy itself . In particular, Hanberger (2002) argues the case for pursuing increased acceptance or improved legitimacy through discursive or deliberative democracy. “To arrive at legitimate solutions to highly complex social problems, considering local conditions and priorities, a participatory or discourse democratic model seems more appropriate.” (Hanberger, 2002, p. 30). Mouffe, Zolo, Gabardi, all argue more generally for an ongoing participatory (generally agonistic) interaction as the only vehicle through which the tensions of modern democratic institutions in a society of diverse perspectives can be accommodated (though never resolved). It is this concern that drives the following section.

Participatory evaluation within representative government?

“Your representative owes you, not his industry only, but his *judgement*; and he betrays, instead of serving you, if he sacrifices it to your *opinion*. (Burke, Edmund, 1774; emphasis added.)

“The most important thing that should not happen is the sidelining of elected officials and legislatures. E-governance holds the promise of greater engagement between governments and citizens, but leaves the role of legislatures increasingly unclear, particularly in a Westminster system that concentrates power in ministers. Initiatives must incorporate elected officials in their roles as representatives, intermediaries and particularly instruments of accountability.” (Oliver and Sanders, 2003, p. 4)

In Edmund Burke’s classic justification for evidence-based decision by informed representatives in a parliamentary democracy, the judgement of the representative trumps the opinion of the citizen presumably because of the logic of specialization and division of labour in part, but perhaps more strongly on the basis of the more informed position from which representatives, as a consequence, may be presumed to act.

The same tradition seems to persist in the presumption that the stronger foundation offered by objective and rigorous ‘sound science’ and factual evidence permits a more informed and more rational decision inside government than can be based on any aggregation of the perceptions and opinions of those outside the elite specialist cadres. It is that premise that the literature mentioned above calls into question, suggesting that there must be a more broadly-based foundation for a more ‘socially robust’ science. There is no longer unquestioning acceptance of the privileged position of scientific method, or the authoritative dominance of formal credentialed expertise.

The concern

In general, given the realities of scale, complexity and limits on reasoning potential – not to mention the reality of ‘consultation fatigue’ as already a serious problem in some communities – arguments for reliance on participatory mechanisms might seem to be a counsel of despair. As noted again in the recent European survey of civil society and participation mechanisms (Banthien, 2003), participation is not the answer

in all cases, and in some cases may be counterproductive. As the Arnstein ladder (Arnstein, 1969) suggests, there are many forms and degrees of citizen engagement, and different approaches to participation will be appropriate in different circumstances. None will be without drawbacks, and a major concern will be the legitimacy and capacity of the civil society organizations through which participation may be channelled. In the proposals for change set out in its 2001 White Paper on European Governance, the European Commission heads its list with the goal of better involvement and more openness, and goes on to a commitment that the Commission will

“establish and publish minimum standards for consultation on EU policy” and “establish partnership arrangements going beyond the minimum standards in selected areas committing the Commission to additional consultation *in return for more guarantees of the openness and representativity of the organizations consulted*” (European Commission, 2001, p. 4, emphasis added).

This concern with the legitimacy of the civil society organizations participating in deliberative fora for policy formation and program evaluation is fundamental, but its exploration is a whole new study.

The central question for the moment seems to be whether it is possible that e-governance, based on the ICT revolution, could offer a way out, could re-establish in some sectors sufficient trust in inclusive guiding processes that people would be willing once again to rely on specialized disciplinary or technical expertise (including local and traditional knowledge and other ways of knowing) in implementation of collective decisions pursuing agreed broad value orientations.

And the particular challenge is to know how this trick can be accomplished while strengthening the institutions of parliamentary democracy, using civil society organizations as representative institutions as well, but recognizing still that

ultimately accountability and compliance must be grounded in the informal life-world, not formal organizations. Legislatures must be accountable for decisions of general application, but individuals are the only entities that can take action to comply with the intent, and realize the purpose, of collective intentions expressed by a legislative body.

Some illustrative possibilities

This paper was intended primarily as a review of the literature and experience in evaluation from which the panel might draw some general lessons. Nonetheless, some specific implications or possibilities might be explored.

The general conclusion from this paper is that the government should move toward setting and undertaking evaluation processes in the context of more open governance, participation and e-democracy, not in the context of 'inside science' and generally accepted accounting principles or performance reporting. Specifically, this shift might be attempted by

- Requiring that, beyond routine or cyclical program monitoring for management purposes, evaluations generally be undertaken in some form of broader setting, perhaps an e-forum, right from the beginning of the planning for evaluation, from the evaluability assessment through to the appraisal of results, with the model of participatory integrated assessment in mind;
- Emphasizing meta-evaluation or systematic review rather than new appraisal (the work of the Campbell Collaboration provides a framework, a model and emerging guidelines for such systematic reviews; see www.campbellcollaboration.org);

- Learning from experience (as might be envisioned within the conceptual framework set out by Kathryn Bouey in remarks on information as a public resource (Bouey, 2003);
- Mandating the launching of such a review periodically for all significant programs for which assessment is desired, with recognition that senior management or Ministers might observe or participate at any time;
- Ensuring that departments and agencies have staff to assist with such fora as well as with public access to relevant information;
- Developing dedicated resources and specialized staff to support MPPs in assisting constituents to find the information they want and to participate in evaluation where they wish; insisting that departments give special priority to MPPs and their staff in this function (see Turner, 2003).
- More speculatively, perhaps, consider all this taking place following modest reform of the legislature to create a legislative presence in ongoing sectoral roundtables, or to extend the functions of standing committees to include such ongoing roundtables. Following the model of the very effective special Parliamentary Task Forces introduced by the Clark administration, imagine legislative or parliamentary standing committees able to claim public service expertise in analytical capacity and access to information, and able to engage stakeholders in inclusive deliberative processes of the kind described above.

Obviously there are no silver bullets, and this paper does not propose any magic solution. It represents, at heart, merely a plea to learn from the experience of the last few decades, to rely on social learning, to keep audit ideas in their place and to avoid the mistake of substituting technical rationales for value-based consensus – and, when thinking about institutions for evaluation activities in the future, to look seriously at what is already going on in attempts to grapple with environmental policy,

sustainable development, and other fields in which the relevance of vernacular understanding, the need for inclusive deliberation and the possibility of socially robust expertise has already been recognized.

Conclusions

“There may be some benefits from initiatives which seek to introduce more instrumental rationality into the policy making process, but *there is even more to be gained from opening up policy making processes: enabling participation by a wide range of stakeholders and citizens.* (Nutley, Sandra. 2003. “Bridging the policy/research divide: Reflections and lessons from the UK. Keynote paper for the National Institute of Governance Conference, Canberra, Australia, 23/24 April, 2003. P. 15, emphasis in original)

“Government cannot fix things because things are never fixed: all is flux and uncertainty....The shift to communicative modes of policy making involves the recognition that although we cannot know, we can learn and the role of government is to facilitate private and public learning.” (Parsons, 2002, p.6) “We must become adept at learning. We must become able not only to transform our institutions, in response to changing systems and requirements; we must invent and develop institutions which are ‘learning systems’, that is to say capable of bringing about their own continuing transformations.” (Schon, 1973 as cited in Parsons, 2002)

The story line for this paper has run as follows.

There are distinct approaches to project evaluation, program evaluation, performance evaluation, and policy evaluation, running from the most ‘factual’ and amenable to uncontested characterization to least ‘factual’ and most subject to conflicting and contested perspectives¹⁰.

Similarly, scholars and students of operations research have long argued the distinction between ‘tame’ or simple problems, and ‘wicked’ or ‘messy’, complex

¹⁰ This spectrum in practice mirrors the spectrum from objectivity to subjectivity in truth claims described by Putnam (1987), who argued forcefully for recognizing a continuum rather than dualism in such claims.

problems. For the former, there are ways to agree reasonably readily on the nature of objective realities, on the extent to which hypotheses are accepted or resisted by those realities ('falsification'), and on the relevance of evidence; issues are generally amenable to technical solutions, and convergence to some workable consensus on relevant evidence is feasible. For the latter problems, none of these things is true.

For the former, where facts, figures of arithmetic, and technical or engineering considerations dominate, accounting and audit can support informative reporting. For the latter, one is left with rhetoric and figures of speech, with only a process of narrative, dialogue, and deliberation to rely upon to achieve convergence within what one hopes will be accepted as a legitimate communicative forum. If it is so accepted, then policy intentions may be realized, and individual consequences or outcomes accepted even when they are adverse or painful.

This paper has not attempted to add to the vast array of commentary and guidance on the conduct of appraisal at the scale of project evaluation. Rather it has reviewed a few of the reasons leading to a belief that most work of any great policy relevance in evaluation must be found at the highly frame-dependent, contested end of the spectrum. It has argued that most serious program evaluation falls toward that end, and can not usefully be approached from the perspective of the comprehensive auditor working to prescribed conventional standards, or the related treasury or management board guidance that relies on that perspective. Their work falls into the trap of 'the modern fact' as an assembly of allegedly incontrovertible assertions that have only precision, not verifiable accuracy, to support them.

Most evaluation that matters must therefore be approached through the realm of interactive deliberative processes, or democratic evaluation. There are many variants,

a few of which have been reviewed briefly in this paper. But fundamentally all are in search of communicative action. The fundamental question is whether any of this can be accomplished at other than local scale where personal engagement and the building of bonding social capital might be feasible.

Another way of expressing the summary argument on which this paper is based is to recall that policy is only defined and realized when compliance with policy intentions is achieved. Assuring compliance in government programs aimed at delivery of services to individuals or individual organizations is the *implementation problem*; it entails achieving the commitment of the wide range of partly autonomous agents and subordinate organizational units whose cooperation is necessary to achieve the intent of the program on a continuing basis, at street level, distant from central authorities.

But for regulatory or hortatory programs, incentive programs, educational programs and other instruments directed toward changing individual awareness, beliefs, values and preferences, the success of policies demands addressing the *compliance problem*, and ultimately the internalization of compliance in fundamental norms and values. It means changing the behaviour of individuals for purposes that are collective in nature, establishing norms that benefit all by influencing social (not judicial) interaction and behaviour.

So this latter approach is inherently collective, addressed to collective action problems and dilemmas. The former is individual, addressed to service to an individual client whose desires and preferences are assumed given, not targets for change. For these the standard machinery of economic science is relevant, and market-shadowing techniques such as cost-benefit analysis or cost-effectiveness analysis are helpful to managers attempting to increase the efficiency and effectiveness of programs fulfilling

given ends. Conventional formative, utilization-centred evaluation is an appropriate tool.

The latter are inherently political, and the criteria are inevitably ethical; for these a different mode of evaluation of government performance is essential, and it is to the considerations around the design of these that this paper has been directed.

More particularly, what should Ontario do? Invest less in efforts to report massive amounts of numbers, perhaps, and reallocate resources to reporting relevant indicators to public fora, commissioning work by networks of NGOs, and supporting ongoing inclusive epistemic communities. Put resources into strengthening the abilities of representative institutions to deal better and more openly with citizens. Consider standing committees or special task forces of the legislature as ongoing sectoral roundtables for inclusive discussion and deliberation, and support these with all the information and communications capacity of the public service. Trust the citizen with more information, and rest more faith in democracy than in modern controllership. Go the route neither of populism nor elitism, but of continuing community. This means pursuing the goal of something like continuing inquiry and democratic deliberative evaluation, exploiting the vision of e-democracy.

It is possible, as argued in the previous section, that institutional arrangements to strengthen representative institutions could be developed in the context of a participatory e-democracy. In such a setting, a deliberative democratic approach to evaluation could be pursued in inclusive processes at various scales, as part of ongoing conversations, ongoing social learning, ongoing adaptive management, in ongoing resilient communities.

All of this takes us back, of course, to government as a learning organization, but more generally to governance processes based on citizen engagement.

As a practical matter, the conclusion is simply to pursue a division of labour, appropriately using very different tools for very different tasks. Deal with the engineering issues – the allocation of available (scarce) resources optimally in the pursuit of agreed goals – in the routine positivist fashion, and deal with the ethical (political) issues through the pursuit of communicative action. Recognize that the apparatus of the former offers no satisfactory route to the latter. Recognize that there is little interest on the part of legislators (or Ministers or citizens) in A-base review or zero-based budgeting, except perhaps for some reassurance that it is happening as part of the housekeeping that is expected to be entrenched in ongoing good management processes.

From yet another perspective, the recommendation is to respond to the public perception of greater risks in a more uncertain complex world by moving toward greater resilience through greater discretion, devolution and reliance on trust, and away from a focus on the competing goal of formal external accountability that compromises achievement of that objective.

The pursuit of equal opportunity for all, for the flourishing of human potential, can only be pursued, it may be argued, through the self-determination made possible within the institutions and norms of a knowledge-based democracy and an experimenting (learning) society. It is this vision, not a model of formal accountability, that should underlie the evolution of evaluation activity in Ontario over the coming decades.

Bibliography

Alkin, Marvin C. 2003. "Evaluation Theory and Practice: Insights and New Directions", *New Directions for Evaluation*, 97, pp81-89.

Arias, Ernesto G. and Gerhard Fischer. 2000. "Boundary Objects: Their Role in Articulating the Task at Hand and Making Information Relevant to It". Boulder, Colorado: University of Colorado, Center for LifeLong Learning & Design and Institute of Cognitive Science, Department of Computer Science, and College of Architecture and Planning.

Armstrong, James L. 1997. "Reason and Passion in Public Sector Reform", a discussion paper prepared for the PSC Learning Series. Ottawa: Public Service Commission of Canada.

Armstrong, James L. 1997. "Stewardship and the Public Service", a discussion paper prepared for the PSC. Ottawa: Public Service Commission of Canada.

Arnstein, Sherry R. 1969. "A Ladder of Citizen Participation", *Journal of the American Planning Association*, 35, 4 (July, 1969) pp 216-224.

Auditor General of Canada. 1990. "Report of the Auditor General of Canada - Chapter 1: Matters of Special Importance and Interest", retrieved May 4 2003, from the Office of the Auditor General of Canada website: <http://www.oag-bvg.gc.ca/domino/reports.nsf/html/9001ce.html>.

Bäckstrand, Karin. 2002. "Civic Science for Sustainability Reframing the Role of Scientific Experts, Policy-makers and Citizens in Environmental Governance", paper prepared for the 2002 Berlin Conference on the Human Dimensions of Global Environmental Change 'Knowledge for Sustainability Transition: The Challenge for Social Science' Berlin December 6-7 2002. Cambridge, Massachusetts: Massachusetts Institute of Technology.

Banthien, Henning et al. 2003. *Governance of the European Research Area: The role of civil society*. IFOK Interim Report, May 16, 2003.

Beck, Ulrich. 1992. *Risk Society: towards a new modernity*. London: Sage

Berkes, Fikret and Carl Folke (eds). 2000. *Linking social and ecological systems : management practices and social mechanisms for building resilience*, edited by Fikret Berkes

and Carl Folke ; and with the editorial assistance of Johan Golding. Cambridge: Cambridge University Press (c. 1998)

Boisclair, Jean-Pierre, 1997. *Report of the Independent Review Panel on Modernization of Comptrollership in the Government of Canada*. Ottawa: The Panel.

Bouey, Kathryn A. 2003. "Information as Public Resource", Breakaway Session, Crossing Boundaries Conference. Ottawa: Government of Ontario.

Bourgogne, Jocelyne. 1998. at http://www.pco-bcp.gc.ca/default.asp?Language=E&Page=clerksspeechesmessages&Sub=clerksspeeches&Doc=1998_symp_e.htm)

Burke, Edmund. 1774. Speech to the Electors of Bristol, 3 November 1774, accessed June 2003 at <http://press-pubs.uchicago.edu/founder/documents/v1ch13s7.html>.

Campbell, D. T. 1969. "Reforms as Experiments", *American Psychologist* 24, pp.409-429

Canada, Privy Council Office. 2003. A Framework for the Application of Precaution in Science-based Decision Making about Risk.

Canada, Treasury Board Secretariat. 2003. "Heads of Evaluation, Community Strategic Planning Session (March 26, 27)". Accessed August 5, 2003 at www.tbs-sct.gc.ca/eval/dev/mar26_2003-PR_e.asp?printable+True

Caracelli, Valerie, J. and Hallie Preskill, eds. 2000. "The Expanding Scope of Evaluation Use", *New Directions for Evaluation*, 88. San Francisco: Jossey-Bass.

Carlsson, Lars. 2000. "Non-Hierarchical Evaluation of Policy", *Evaluation*, 6(2), pp201-216.

Cash, David and Karen Fisher-Vanden, et al. 1997. "Chapter IV - Working Group 3 Background Paper: How can assessment processes and outcomes be improved?" retrieved November 26 2000, from Institute of Global Environment and Society website: <http://grads.iges.org/geaproject1997/gea5.html>.

Cash, David *et al.* 2002. "The Institutional Challenges of Producing Salient, Credible and Legitimate Regional Integrated Assessments" International Workshop on Regional Integrated Assessment of Climate Impacts, Harvard, Sept 16-20, 2002.

Casti, John L. 1994. *Complexification: Explaining a paradoxical world through the science of surprise*, New York: Harper Collins.

CCAF-FCVI Inc. 2002. *Reporting Principles: Taking Public Performance Reporting to a New Level*. Ottawa: CCAF-FCVI Inc.

Centre of Excellence for Evaluation, Comptrollership Branch, Treasury Board Secretariat. 2001. "Guide for the Development of Results-based Management and Accountability Frameworks". Ottawa: Treasury Board of Canada Secretariat.

Centre of Excellence for Evaluation, Comptrollership Branch, Treasury Board Secretariat. 2002. "Guidance for Strategic Approach to Results-based Management and Accountability Frameworks". Ottawa: Treasury Board of Canada Secretariat.

Centre of Excellence for Evaluation, Comptrollership Branch, Treasury Board Secretariat. 2002. "Companion Guide: The Development of Results-based Management and Accountability Frameworks for Horizontal Initiatives". Ottawa: Treasury Board of Canada Secretariat.

Chrisman, Nicholas. 1999. "Trading Zones or Boundary Objects: Understanding Incomplete Translations of Technical Expertise", presentation retrieved July 12 2003, from the University of Washington Faculty Web Server:
<http://faculty.washington.edu/chrisman/Present/4S99.html>.

Christie, Christina A. 2003. "Editor's Notes: The Practice-Theory Relationship", *New Directions for Evaluation*, 97, pp1-5.

Christie, Christina A. 2003. "What Guides Evaluation? A Study of How Evaluation Practice Maps onto Evaluation Theory", *New Directions for Evaluation*, 97, pp7-35.

Christie, Christina A. 2003. "Understanding Evaluation Theory and Its Role in Guiding Practice: Formal, Folk, and Otherwise", *New Directions for Evaluation*, 97, pp91-93.

Clark, William, et al. 2002. "Information as Influence: How institutions mediate the impact of scientific assessments on global environmental affairs". Cambridge, Massachusetts: John F. Kennedy School of Government, Harvard University, Faculty Research Working Papers Series.

Clark, William and Nancy Dickson. 1999. "The Global Environment Assessment Project: Learning from Efforts to Link Science and Policy in an Interdependent World",

retrieved May 21 2003, from US Global Change Research Program website:

<http://earth.usgcrp.gov/usgcrp/Library/nationalassessment/newsletter/1999.10/GEA.html>.

Coglianesse, Cary. 2002. "Is Satisfaction Success? Evaluating Public Participation in Regulatory Policymaking". Regulatory Policy Program Working Paper RPP-2002-09. Cambridge, MA: Center for Business and Government, John F. Kennedy School of Government, Harvard University.

Cappe, Mel. 2000. at http://www.pco-bcp.gc.ca/default.asp?Language=E&Page=clerksspeechesmessages&Sub=clerksspeeches&Doc=2000_sustain-dev_e.htm

Coglianesse, Cary. 2002b. "Empirical Analysis and Administrative Law", Regulatory Policy Program Working Paper RPP-2002-10. Cambridge, MA: Center for Business and Government, John F. Kennedy School of Government, Harvard University.

Coglianesse, Cary. 2003. "The Internet and Public Participation in Rule-Making", Regulatory Policy Program Working Paper RPP-2003-05. Cambridge, MA: Center for Business and Government, John F. Kennedy School of Government, Harvard University.

Commission of the European Communities, 2001. "Democratising Expertise and Establishing Scientific Reference Systems: Report of Working Group 1b". Brussels.

Culpepper, Pepper. 2003. "Institutional Rules, Social Capacity and the Stuff of Politics: Experiments in Collaborative Governance in France and Italy". Working Paper RWP03-029, John F. Kennedy School of Government, Harvard University.

Culver, Keith C. "Making Public/Private Research Partnerships Work: Finding the Right Road to E-Democracy". Fredericton: University of New Brunswick, Centre for Social Innovation Research.

Culver, Keith C. "Innovation is Not Reform: Can Democracy Survive the New ICTs". Fredericton: University of New Brunswick, Centre for Social Innovation Research.

Culver, Keith C. "Embracing Innovations: E-Democracy in New Brunswick, Canada". Fredericton: University of New Brunswick, Centre for Social Innovation Research.

Cummings, Rick. 2002. "Rethinking Evaluation Use", paper presented at the 2002 Australasian Evaluation Society International Conference. Perth, Western Australia: Murdoch University.

Datta, Lois-ellin. 2003. "Important Questions, Intriguing Method, Incomplete Answers", *New Directions for Evaluation*, 97, pp37-46.

de Jong, Simone, Petra Ahrweiler and Paul Windrum. 2001. "Theoretical Considerations of the SEIN Evaluation Approach S.E.A. I", SEIN-Project Paper No. 15, The SEIN-Project. The Netherlands: University of Maastricht.

Department of Finance Canada. 2003. "Building the Canada We Want, Budget 2003: Improving Expenditure Management and Accountability". Ottawa: Her Majesty the Queen in Right of Canada.

Dobell, Rodney and David Zussman. 1981. "If politics is theatre then evaluation is (mostly) art", *Canadian Public Administration*, 24(3), pp404-427.

Dobell, Rodney. 1989. "Fresh Thoughts in Fat City: Innovation in the Federal Public Service", Keynote address to APEX Innovation Symposium, Ottawa, January 19, 1989.

Dobell, Rodney. 1997. "Complexity, Connectedness and Civil Purpose: Public Administration in the Congested Global Village", *Canadian Public Administration* 40, 2, pp 346-370.

Dobell, Rodney, 1997. (with Luc Bernier) "Citizen-centered governance: Implications for inter-governmental Canada" in *Alternative Service Delivery*. Robin Ford and David Zussman (eds.). IPAC-KPMG, 1997.

Dobell, Rodney. 1999. "Evaluation and Entitlements: Hartle's search for rationality in government", in Bird, Trebilcock and Wilson, eds, *Rationality in Public Policy: Retrospect and Prospect, A Tribute to Douglas G. Hartle*. Toronto: Canadian Tax Foundation.

Dobell, Rodney. 2000. *Social Risk, Political Rationality and Official Responsibility: Risk Management in Context*. Study prepared for the Walkerton Inquiry.

Douglas, Mary. 1999. "Four cultures: the evolution of a parsimonious model", *Geojournal* 47, pp. 411-415.

Dubnick, Melvin, J. 1999. "Demons, Spirits, and Elephants: Reflections on the Failure of Public Administration Theory", prepared for delivery at the 1999 Annual Meeting of the American Political Science Association.

Duhamel, Ron, M.P. 1996. "IRPP Evaluation Report – Parliamentary Working Group", retrieved May 4 2003, from the Treasury Board of Canada Secretariat website: http://www.tbs-sct.gc.ca/rma/eppi-ibdrp/irp-arp/pwg-gtp_e.asp.

Duignan, Paul. 2001. "Upping the Evaluation Debate: What We Need to Do Next", paper presented at the Measuring Public Sector Performance 2001 Conference, Wellington, New Zealand. Auckland, New Zealand: Parker Duignan Ltd.

Eden, Sally. 1996. "Public participation in environmental policy: considering scientific, counter-scientific and non-scientific contributions", *Public Understanding of Science*, 5, 183-204.

European Commission. 2001. European Governance: A White Paper. Com(2001) 428 final. Brussels: Commission of the European Communities.

Evans, Gord and Nick Manning. "A practical approach to assessing central government policy-making institutions in cabinet government: learning lessons from recent World Bank institutional analyses", preliminary draft.

Fetterman, David. 2003. "Fetterman-House: A Process Use Distinction and a Theory", *New Directions for Evaluation*, 97, pp47-52.

Floc'hlay, Beatrice and Eric Plottu. 1998. "Democratic Evaluation: From Empowerment Evaluation to Public Decision-Making", *Evaluation*, 4(3), pp261-277.

Forester, John. 1999. *The Deliberative Practitioner Encouraging Participatory Planning Processes*. Cambridge, Massachusetts: The MIT Press.

Forss, Kim and Knut Samset. 1999. "Square Pegs and Round Holes", *Evaluation*, 5(4), pp407-421.

French, Richard. 1992. "Postmodern Government", *Optimum* 23, 1. pp 42-51.

Furubo, Jan-Eric. 2003. "The Role of Evaluation in Political and Administrative Learning and the Role of Learning in Evaluation Praxis", paper prepared for the Learning Government Symposium, OECD, 3-4 February, 2003.

- Geva-May, I and Pal, L. 1999. "Policy evaluation and analysis: Exploring the differences", *Policy Evaluation* 5, 1, pp. 5-18
- Gieryn, T. T. 1983. "Boundary work and the demarcation of science from non-science", *American Sociological Review*, 48, pp 781-795.
- Goldoff, Anna C., ed. "Symposium on Chaos Theory and Management", retrieved June 29, 2003 from the Public Administration and Management: An Interactive Journal website: <http://www.pamij.com/goldoff.html>.
- Greene, Jennifer C. 2002. "Towards Evaluation as a 'Public Craft' and Evaluators as Stewards of the Public Good or On Listening Well", keynote address presented at the 2002 Australasian Evaluation Society International Conference. Urbana-Champaign: University of Illinois.
- Gregory, Amanda. 2000. "Problematizing Participation: A Critical Review of Approaches to Participation in Evaluation Theory", *Evaluation*, 6(2), pp179-199.
- Guba, Egon G. and Yvonna S. Lincoln. 1989. *Fourth Generation Evaluation*. London: Sage.
- Guba, Egon G. and Yvonna S. Lincoln. 2001. "Guidelines and Checklists for Constructivist (a.k.a. Fourth Generation) Evaluation". The Evaluation Centre, Kalamazoo: Western Michigan University.
- Gunderson, Lance et al (eds). 1995. *Barriers and Bridges to the Renewal of Ecosystems and Institutions*. New York: Columbia University Press.
- Guston, David H., et al. 2000. "Report of the Workshop on Boundary Organizations in Environmental Policy and Science". Cambridge, Massachusetts: John F. Kennedy School of Government, Harvard University, Belfer Center for Science & International Affairs.
- Hanberger, Anders. 2002. "Democratic implications of public organizations: A Weberian methodology to assess organizational implications for democracy and legitimacy", paper to be presented at "Nordiska statsvetenskapskongressen" (NOPSA), August 2002. Umeå: Umeå Centre for Evaluation Research, Umeå University.
- Henry, Gary T. and Melvin M. Mark. 2003. "Toward an Agenda for Research on Evaluation", *New Directions for Evaluation*, 97, pp69-80.

HM Government. 2002. "In the service of democracy: a consultation paper on a policy for electronic democracy", UKonline.

HM Treasury. 1997. *Appraisal and Evaluation in Central Government*. London: Her Majesty's Stationery Office.

HM Treasury. 2003. *The Green Book: Appraisal and Evaluation in Central Government*. London: TSO.

House, Ernest R. and Kenneth R. Howe. 1998. "Deliberative Democratic Evaluation", in Ryan and DeStefano (eds) *Evaluation as a Democratic Process: Promoting Inclusion, Dialogue, and Deliberation*. *New Directions in Evaluation* 85 pp. 3-12.

House, Ernest R. 2000. "The Limits of Cost Benefit Evaluation", *Evaluation*, 6(1), pp79-86.

House, Ernest R. 2001. "Responsive Evaluation (and Its Influence on Deliberative Democratic Evaluation)", *New Directions for Evaluation*, 92, pp23-30.

House, Ernest R. and Kenneth R. Howe. 2000. "Deliberative Democratic Evaluation Checklist", retrieved May 19 2003, from The Evaluation Centre, Evaluation Checklists website: <http://www.wmich.edu/evalctr/checklists/>.

Human Resources Development Canada. 1998. "Quasi-Experimental Evaluation", Ottawa: Evaluation and Data Development, Strategic Policy, HRDC.

Human Resources Development Canada. 1998. "The Use of Social Indicators as Evaluation Instruments - Final Report", Ottawa: Strategic Evaluation and Monitoring, HRDC.

Kates, Robert W. and William C. Clark, et al. 2002. "Sustainability Science". Cambridge, Massachusetts: John F. Kennedy School of Government, Harvard University, Faculty Research Working Papers Series.

Kickert, W. et al. 1997. *Managing Complex Networks*. London: Sage.

King, Jean A. 2003. "The Challenge of Studying Evaluation Theory", *New Directions for Evaluation*, 97, pp57-67.

Kirkhart, Karen E. 2000. "Reconceptualizing Evaluation Use: An Integrated Theory of Influence", *New Directions for Evaluation*, 88, pp5-23.

Julnes, Patria de Lancer and Marc Holzer. 2001. "Promoting the Utilization of Performance Measures in Public Organizations: An Empirical Study of Factors Affecting Adoption and Implementation", *Public Administration Review*, 61(6), pp693-705.

Lee, Linda E. and Bud Long. 2002. "Definition of the Core Body of Knowledge in Evaluation: the Situation", retrieved November 23 2002, from The Canadian Evaluation Society website: <http://www.evaluationcanada.ca/>.

Lenihan, Donald G. 2002. "Realigning Governance: From E-Government to E-Democracy", Ottawa: The Centre for Collaborative Government.

Lenihan, Donald G. 2002. "E-Government, Federalism and Democracy: The New Governance", *Changing Government*, 9. Ottawa: The Centre for Collaborative Government.

Lenihan, Donald G. and Tony Valeri. 2003. "Finding our Digital Voice: Governing in the Information Age - The Crossing Boundaries Final Report and Recommendations", *Changing Government*, 11. Ottawa: The Centre for Collaborative Government.

Light, Richard J. 2001. "Evaluating to Resolve Controversies", *New Directions for Evaluation*, 90, pp3-8.

Lindblom, Charles E. and David K. Cohen. 1979. *Usable Knowledge*. New Haven: Yale University Press.

Lynn, Jr., Laurence E. 2001. "Public Management", *Handbook of Public Administration*. San Francisco: Sage Publications.

Lynn, Jr., Laurence E., Carolyn Heinrich and Carolyn Hill. 2003. "A Logic of Governance: A Framework for Studying the Effectiveness of Public Policy and Programs", *Research Summary*, 1, 11. Chicago: University of Chicago Press.

Muller, Michael J. 1999. "Translation in HCL: Toward a Research Agenda". Submitted to *Transactions on Computer-Human Interaction*. Cambridge, Massachusetts, Lotus Development Corporation.

Muller-Clemm, W. J. and Barnes, M. P. 1997. "A historical perspective on federal program evaluation in Canada", *The Canadian Journal of Program Evaluation* 12, 1 pp. 47-70

Nagy, Magdolna Toth, 2002. "Improving Interaction with Citizens". Paris: OECD.

Nowotny, Helga, Peter Scott and Michael Gibbons. 2003. "Re-thinking Science: Mode 2 in Social Context". Forthcoming in *Minerva* (2003), retrieved July 16, 2003 from the Dr. Helga Nowotny website:
http://www.nowotny.ethz.ch/pdf/Nowotny_Gibbons_Scott_Mode2.pdf.

OECD Public Management Policy Brief No. 5. 1998. "Best Practice Guidelines for Evaluation".

OECD Public Management Service. 1999. *Improving Evaluation Practices*. PUMA/PAC(99)1. Paris: OECD.

OECD Public Management Policy Brief No. 10. 2001. "Engaging Citizens in Policy-making: Information, Consultation and Public Participation". Paris: OECD

OECD. 2001. *Citizens as Partners: OECD Handbook on Information, Consultation and Public Participation in Policy-making*. Paris: OECD Publications Service.

OECD Public Management Service. 2002. "Performance Management Program in the Canadian Federal Public Service", paper for OECD-Germany High-Level Symposium, Berlin.

OECD. 2002. *Distributed Public Governance*. Paris: OECD.

Office of the Auditor General of British Columbia. 2003. *Building Better Reports: Our Review of the 2001/02 Reports of Government*. Victoria: Queen's Printer for British Columbia.

Oliver, E. Lynn and Larry Sanders. 2003. "The Regina Declaration: E-government for the 21st Century", *Optimum Online* 33(2), retrieved July 6, 2003 from Optimum Online The Journal of Public Sector Management website:
www.optimumonline.ca/print.phtml?id+177

Online NewsHour. 2003. "By the People: A National Conversation about America in the World". MacNeil/Lehrer Productions.

Ontario, Cabinet Office. April, 2000. *Transforming Public Service for the 21st Century*.

Ontario, Management Board Secretariat. March, 2003. *Guide to Program Evaluation: Work in Progress*. Toronto: Queen's Printer of Ontario.

Paquet, Gilles. 1999. *Governance Through Social Learning*. Ottawa: University of Ottawa Press.

Patton, Michael Quinn. 1978. *Utilization-focused Evaluation*. London: Sage

Patton, Michael Quinn. 1997. *Utilization-focused Evaluation: The New Century Text*. (3rd ed). London: Sage.

Patton, Michael Quinn. 1998. "Discovering Process Use", *Evaluation*, 4(2), pp225-233.

Patton, Michael Quinn. 2002. "Utilization-Focused Evaluation (U-FE) Checklist", retrieved May 19 2003, from The Evaluation Centre, Evaluation Checklists website: <http://www.wmich.edu/evalctr/checklists/>.

Pawson, Ray and Nick Tilley. 1998. "Cook-Book Methods and Disastrous Recipes: A Rejoinder to Farrington", *Evaluation*, 4(2), pp211-213.

Pawson, Ray. 2002. "Evidence-based Policy: The Promise of 'Realist Synthesis'", *Evaluation*, 8(3), pp340-358

Pellizzoni, Luigi. 2001. "The myth of the best argument: power, deliberation and reason", *British Journal of Sociology*, 52(1), pp59-86.

Perrin, Burt. 2002. "Implementing The Vision: Addressing Challenges to Results-Focused Management and Budgeting", paper based upon a presentation to the OECD meeting 'Implementation Challenges in Results Focused Management and Budgeting' 11-12 February 2002, Paris. Paris: OECD Publications Service.

Phillips, Susan D. and Michael Orsini. 2002. "Mapping the Links: Citizen Involvement in Policy Processes", CPRN Discussion Paper No. F\21. Ottawa: Canadian Policy Research Networks.

Pollitt, Christopher. 1998. "Evaluation in Europe: Boom or Bubble?", *Evaluation*, 4(2), pp214-224.

- Poovey, Mary. 1998. *A History of the Modern Fact: Problems of Knowledge in the Sciences of Wealth and Society*. Chicago: University of Chicago Press.
- Preskill, H and R.T. Torres. 2000. "The Learning Dimension of Evaluation Use", *New Directions in Evaluation*, 88, pp
- Privy Council Office. 2001. "Guide to Making Federal Acts and Regulations, 2nd ed." Ottawa: Her Majesty the Queen in Right of Canada.
- Putnam, Hilary. 1987. *The Many Faces of Realism: The 1985 Paul Carus Lectures*. LaSalle, Illinois: Open Court.
- Putnam, Hilary. 2002. *The Collapse of the Fact-Value Dichotomy*. Cambridge, Massachusetts: Harvard University Press.
- Queensland Government. 2001. "E-democracy policy framework". Queensland, Australia: E-Democracy Unit.
- Ramlow, Mary E. 2003. "The Program Evaluation Standards: Summary of the Standards", retrieved May 19 2003, from The Evaluation Centre, Evaluation Checklists website: <http://www.wmich.edu/evalctr/checklists/>.
- Robertson, G. 1838. "Exclusion of Opinions", *London and Westminster Review*, 61 (April, 1838, p. 37)
- Rogers, Patricia J., Timothy A. Hacsí, Anthony Petrosino and Tracy A. Huebner, eds. 2000. "Program Theory in Evaluation: Challenges and Opportunities", *New Directions for Evaluation*, 87. San Francisco: Jossey-Bass.
- Ryan, Katherine E. and Lizanne DeStefano, eds. 2000. "Evaluation as a Democratic Process: Promoting Inclusion, Dialogue, and Deliberation", *New Directions for Evaluation*, 85. San Francisco: Jossey-Bass.
- Salomon, Jean-Jacques. 2001. "Society talks back", *Nature*, 412, pp585-586.
- Sanderson, Ian. 2000. "Evaluation in Complex Policy Systems", *Evaluation*, 6(4), pp433-454.
- Sanderson, Ian. 2002. "Making Sense of 'What Works': Evidence-Based Policy Making as Instrumental Rationality?", paper presented at the Political Studies

Association Annual Conference Aberdeen, 5th – 7th April 2002. Leeds: Leeds Metropolitan University, Policy Research Institute.

Schon, Donald A. 1971. *Beyond the Stable State*. London: Temple Smith.

Schon, Donald A. and Martin Rein. 1994. *Frame Reflection*. New York, NY: Basic Books.

Scott, James. 1998. *Seeing Like a State*. New Haven: Yale University Press.

Scriven, Michael. 2002. "Key Evaluation Checklist", retrieved May 19 2003, from The Evaluation Centre, Evaluation Checklists website:
<http://www.wmich.edu/evalctr/checklists/>.

Segsworth, Bob. 2003. "Accountability, Evaluation and Performance Monitoring: A Comparative Perspective. Toronto: Panel on the Role of Government Research Paper Series RP(15), April 2003.

Sen, Amartya. 1987. *On Ethics and Economics*. Oxford: Basil Blackwell

Social Learning Group. 2001. *Learning to Manage Global Environmental Risks*. Cambridge, Mass: MIT Press.

St. Clair, Asuncion Lera. n.d. "Expert Knowledge and International Poverty Policy: Boundary Work". www.svf.uib.no/svfa/conferences/papers/stclair.doc

Star, Leigh and J.R.Griesemer. 1989. "Institutional ecology, translations and boundary objects". *Social Studies of Science*, 3. pp 387-420.

Stein, Janice Gross. 2001. *The Cult of Efficiency*. Toronto: Anansi.

Stern, P.C. and H.V. Fineberg, eds. 1996. *Understanding Risk: Informing Decisions in a Democratic Society*. Washington: National Academy Press.

Stockton, Margaret. 1994. "Effectiveness Reporting - Guidelines for the Application of the Effectiveness Reporting Framework". Sydney, Australia: Management Improvement Branch, Public Sector Management Office.

Stufflebeam, Daniel L. 2001. "Evaluation Models", *New Directions for Evaluation*, 89. San Francisco: Jossey-Bass.

Taruskin, Richard. 1995. *Text and act : essays on music and performance*. New York: Oxford University Press.

Thagard, Paul. 2001. "Being Interdisciplinary: Trading Zones in Cognitive Science". Waterloo, Ontario: University of Waterloo, Philosophy Department.

The World Bank Group. 2003. "Indicators of Governance and Institutional Quality", retrieved January 15 2003, from The World Bank Group, Governance and Public Sector Reform website: <http://www1.worldbank.org/publicsector/index.cfm>.

Toth, Ferenc, 2001. *Participatory integrated assessment methods*. EEA Technical Report 64. Copenhagen: European Environment Agency, August 2001.

Treasury Board of Canada Secretariat. 2001. "CEE Strategic Plan", retrieved May 4 2003, from the Treasury Board of Canada Secretariat website: http://www.tbs-sct.gc.ca/eval/plan_e.asp.

Treasury Board of Canada Secretariat. 2001. "Study of the Evaluation Function in the Federal Government", retrieved May 4 2003, from the Treasury Board of Canada Secretariat website: http://www.tbs-sct.gc.ca/eval/stud_etud/function-download_e.html.

Treasury Board of Canada Secretariat. 2001. "Results for Canadians: A Management Framework for the Government of Canada", retrieved May 3 2003, from the Treasury Board of Canada Secretariat website: http://www.tbs-sct.gc.ca/res_can/rc_1_e.asp.

Treasury Board of Canada Secretariat. 2001. "Evaluation Policy", Ottawa: Treasury Board of Canada Secretariat.

Treasury Board of Canada Secretariat. 2001. *Program Evaluation Methods: Measurement and Attribution of Program Results*, 3rd ed. Ottawa: Minister of Public Works and Government Services.

Turner, Michael. 2003. Powerpoint presentation prepared for the Crossing Boundaries Conference, Ottawa, May 8, 2003.

Wadsworth, Yoland. 2001. "Becoming Responsive - and Some Consequences for Evaluation as Dialogue Across Distance", *New Directions for Evaluation*, 92, pp45-58.

Waldrop, M. M. 1992. *Complexity: The emerging science at the edge of order and chaos*. New York: Simon and Schuster.

Weiss, Carol H. 1972. *Evaluation Research: Methods of Assessing Program Effectiveness*. Englewood Cliffs, New Jersey: Prentice-Hall Inc.

Weiss, Carol H. 1998. "Have We Learned Anything New About the Use of Evaluation?", *American Journal of Evaluation*, 19(1), pp21-33.

Weiss, Carol H. 1999. "The Interface between Evaluation and Public Policy", *Evaluation*, 5(4), pp468-486.

Wholey, Joseph S. and John W. Scanlon, et al. 1975. *Federal Evaluation Policy: Analyzing the Effects of Public Programs*. Washington: The Urban Institute.

Windrum, Paul and Simone de Jong. 2000. "Foundations of Evaluation Theory: A review", Project Paper No. 9, The SEIN-Project. The Netherlands: University of Maastricht.

Zorzi, Rochelle, Martha McGuire and Burt Perrin. 2002. "Canadian Evaluation Society Project in Support of Advocacy and Professional Development: Evaluation Benefits, Outputs, and Knowledge Elements". Ottawa: Canadian Evaluation Society.