# The Case for Participatory Evaluation in an Era of Accountability

## Jill Anne Chouinard[1]

### Abstract

Evaluation occurs within a specific context and is influenced by the economic, political, historical, and social forces that shape that context. The culture of evaluation is thus very much embedded in the culture of accountability that currently prevails in public sector institutions, policies, and program. As such, our understanding of the reception and use of participatory approaches to evaluation must include an understanding of the practices of new public management and of the concomitant call for accountability and performance measurement standards that currently prevail. In this article, the author discusses how accountability has been defined and understood in the context of government policies, programs, and services and provides a brief discussion of participatory and collaborative approaches to evaluation and the interrelationship between participatory evaluation and technical approaches to evaluation. The main part of the article is a critical look at key tensions between participatory and technocratic approaches to evaluation. The article concludes with a focus on the epistemological and cultural implications of the current culture of public accountability.

## Introduction

Evaluators live in an increasingly multicultural and diverse world, where the need to design evaluations that are culturally responsive to community and that are inclusive of relevant program stakeholders, is growing. While approaches to evaluation have evolved over the years, away from an overreliance on positivist models based on experimental and quasi-experimental methods and toward approaches that are more inclusive and sensitive to local contexts, questions about methodologies and method choice continue to generate significant debate, remaining what Smith (2008) refers to as one of the "fundamental issues" in evaluation. Despite the rich and varied methodological options available, evaluators are nonetheless faced with the challenge of designing and conducting evaluations amid a governance climate where the current gold standard of program evaluation is defined as impartial, objective, and

[1]University of Ottawa, Ottawa, Ontario, Canada

**Corresponding Author:**
Jill Anne Chouinard, University of Ottawa, 72 Harmer Ave, Ottawa, Ontario, Canada K1Y 0T8.
Email: chouinardjill@gmail.com

evidence based (Greene, 2005). The contrast between approaches that are more sensitive and responsive to community needs, such as participatory or collaborative approaches to evaluation, and accountability-driven technocratic approaches, is stark. While multiple approaches to evaluation are now considered methodologically credible within federal public sector contexts, the prevalent vision of evaluation is nonetheless defined as a neutral instrument providing impartial, evidence-based, and objective information intended primarily to satisfy accountability requirements. As such, this approach to evaluation may well fall short of capturing the range of local views, contextualized meanings, and culturally relevant perspectives that are increasingly relevant today, and that participatory and collaborative approaches to evaluations are intended to capture.

Evaluation, as it is envisioned in the current governance context, is intended to serve primarily an instrumental role, as a management tool designed for accountability and decision-making purposes. Evaluation so defined, is thus valued for its perceived scientific and technical authority (House, 1993) and as a way to legitimize government activities, ensure cost-effectiveness, and enhance managerial decision making. Participatory approaches to evaluation, on the other hand, are designed to address diverse program and organizational needs across a broad range of local, program, and cultural contexts, with varied rationales that include local ownership, empowerment, use of findings, organizational and individual learning, and program improvement (Cousins & Chouinard, 2012). In discussing the tension between these two approaches to evaluation, between what is essentially a technical and instrumental vision of evaluation and a contextualized and more democratic participatory approach, this article makes the case for participatory evaluation in the current federal climate of accountability.

To further explore these issues, this article is organized into five key sections. I begin with a review of accountability, tracing its roots to the political and economic exigencies of new public management (NPM). Second, I explore the connections between accountability and evaluation in both Canadian and American contexts, focusing primarily upon federal legislation and policy impacts on evaluation. In section three, I provide a brief overview of participatory and collaborative approaches to evaluation, with a focus on the diverse rationales and contexts of use. I next compare participatory and technological approaches to evaluation through the identification of seven key tensions (e.g., relational, epistemological, pedagogical, contextual, political, methodological, and organizational). In the remainder of the article (section five), I argue that we need to broaden the practice of evaluation, and to seek a place where evaluation theories are not negotiated beyond recognition, and where practice reflects the diversity of our increasingly multicultural society.

## Unpacking Accountability: A Brief Background History

As a pervasive feature of democratic liberal governments, there is no shortage of academic and popular texts on NPM broadly described as a private sector approach to governance characterized by demands for accountability and performance measures in public institutions (Blalock, 1999; Davies, 1999; McDavid & Huse, 2006). Behn (2001) describes NPM as '' the entire collection of tactics and strategies that seek to enhance the performance of the public sector—to improve the ability of governmental agencies *and* their nonprofit and for-profit collaborators to produce results'' (p. 26). While NPM differs slightly in each country, in Canada, NPM has been adopted in the Treasury Board of Canada's results-based agenda and enshrined in the *Management Accountability Framework* (1997) and the *Federal Accountability Act* (2006), policies designed to provide program efficiency and effectiveness and ensure wholesale accountability. In 2009, the Canadian government revised its policies and guidelines on evaluation, articulating a clear vision of evaluation as an evidence-based, neutral, and objective instrument designed to provide outcome and results-based information on program performance and spending. In the United States, principles of NPM are maintained in the *Government Performance and Results Modernization Act* (2010), introduced by

the Obama government to focus on performance information, results, and cost efficiency to manage federal programs. The ethos of NPM is thus based on the principles of economics, efficiency, and effectiveness, with a specific emphasis on accountability for results. While evaluation and performance measurement have each become strategic components of NPM, the focus on accountability and results has led to a more prominent role for performance measurement and a more circumscribed role for evaluation within public sector reform (Mayne & Rist, 2006; McDavid & Huse, 2006). In many ways, program evaluation has thus become a key mechanism of the accountability movement (Chaytor, MacDonald & Melvin, 2002; Norris & Kushner, 2007; Schwandt, 2009), and evaluators the key agents of public accountability (Greene, 1999; Stake, 2001).

While few would likely dispute the need for accountability within our public institutions, the critical question focuses on what accountability actually means (and what it has come to mean) within the context of current Western democratic governance practices. Rarely do people question the need for accountability, nor question how it is being pursued and what is being gained (or lost) through its attainment. As Charlton (2002) argues, "accountability is assumed to be an intrinsically desirable goal, and nobody ever claims that one can have 'too much' accountability-the pressure is always for more" (p. 3). But more of what remains vague, as the mere questioning of accountability in today's governance climate is seen as highly irresponsible. As Cronbach et al. (1980) observed more than 30 years ago, "left undefined, and with its consequences for the system unexamined, accountability is perceived as a good. Who could oppose it?" (p. 133). Thus, while there continues to be significant discussion within the academic milieu, the current "discourse of accountability" circumscribes *public dialogue* about accountability, what it means, who should be involved in its definition, who it is intended for, and what purpose it is intended to serve.

Accountability, according to Charlton (2002), has become a "slippery rhetorical term" (p. 3) with two rather distinct and interchangeable meanings, a technical managerial meaning and a "popular," more general meaning related to democratic responsibility. The conflation of these two conflicting conceptions facilitates the confusion about what we intend and what accountability actually means, as we seamlessly and unconsciously move from a democratic to a technocratic conception. In practice, however, accountability has lost its traditional democratic meaning of shared responsibility and shifted to a predominantly technocratic conception focused on control, regulation, and compliance (Greene, 1999; Perrin, 2007). Not surprisingly, accountability as it is understood today shares many similar traits with Frederick Taylor's scientific management paradigm of the early 20th century, with its ardent focus on management control, efficiency, results, and the objective measurement of predefined and predetermined goals and objectives (Campbell, Fleming, Newell, & Bennion, 1987; Eisner, 1985). Applied to public sector governance, scientific management thus translates into a form of institutional accountability used to "legitimize" (Meyer & Rowan, 1983) government activities and related program spending. Accountability thus becomes what Power (1991) refers to as "an organizational ritual, a dramaturgical performance" (p. 141), what Meyer and Rowan (1983) might have referred to as "myth and ceremony," ultimately leading us to a "culture of compliance" (Franklin, 1990), as we uncritically engage in the increasingly detailed rituals of accountability.

Constructed as a technological, rational concept, accountability has come to assume an apolitical dimension (Biesta, 2004) intended merely to ensure efficiency and effectiveness in government programs and services. In practice, however, accountability is anything but apolitical, as it moves beyond the mere provision of a technical solution to a redesigned system of governance that arguably threatens our democratic principles (Behn, 1998; Power, 1997). The performance management and evaluation architecture of the Canadian government are illustrative of central control and compliance, as government departments are required to complete detailed performance measurement and evaluation plans to satisfy federal accountability requirements. This level of accountability and performance monitoring of public programs and expenditures comes at a cost, as our relationship to the state becomes economic rather than political, thus repositioning us as consumers rather than as

citizens (Biesta, 2004; Ryan, 2007). Noteworthy also are changes between people and professionals, as interactions become increasingly more focused on compliance with performance monitoring, and less on relationships with each other (O'Neill, 2002). Trust in democratic institutions and in relationships between each other and with the state becomes compromised, as it becomes ever more vested in abstract capacities outside of our control, rather than in people (Power, 1997).

## Locating the Intersection of Evaluation and Government Policy Initiatives

Despite the "uneasy peace" (Mark, 2003, p. 191) achieved in the paradigm wars that overshadowed the social sciences in the late 20th century, debates about methodology and method selection continue to dominate the evaluation landscape. The focus of the debate has shifted over the years, from an emphasis on the epistemological divide between qualitative (constructivist) and quantitative (positivist) methodologies, to disagreements about evaluation purpose, the role of the evaluator, the privileging of theoretical models over others, method choice, and the nature of credible evidence and establishing causal connections. While the complexity of this debate likely precludes an easy truce among conflicting sides, it is important to note that despite the profusion of qualitative and mixed methods approaches to evaluation advanced over the past 25 years, the federal policy and program context in many Western democracies nonetheless continues to privilege "scientific" approaches to evaluation (Fitzpatrick, Christie, & Mark, 2009). As Smith (2008) observes, at issue is not whether one methodological approach or one specific method is preferable to another, but the fact that federal governments express a *preference for one approach* over another. While there are undoubtedly exceptions, the following examples from the United States and Canada serve to underscore the general point being advanced that federal departments privilege specific methodological approaches to research and evaluation, at the expense of what are considered less "scientific" approaches.

Although the most recent policy and position papers on program evaluation and design in the U.S. Government Accountability Office (GAO) and in the Office of Management and Budget (OMB) include multiple approaches to evaluation, they nonetheless focus on performance management, impact and outcome evaluation, and experimental and quasi-experimental approaches, all at the expense of other less "rigorous" and "scientific" approaches (GAO, 2012; OMB, 2012). Moreover, despite the fact that the Obama administration has dropped the Program Assessment Rating Tool (a program administered through the OMB that required an assessment of the efficiency and effectiveness of all federal programs, and that considered randomized controlled trials [RCTs] the "gold standard" in program measurement and evaluation), other federal departments nonetheless continue to advocate for such specific approaches to evaluation. For example, in late 2003, the U.S. Department of Education's (DOE) Institute of Educational Sciences issued a public statement clearly articulating a preference for "scientifically based" evaluation methods that give priority to the use of experimental and quasi-experimental designs. According to the DOE Notice (2005):

> . . . the Secretary considers random assignment and quasi-experimental designs to be the most rigorous methods to address the question of project effectiveness . . . proposed evaluation strategies that use neither experimental designs using a matched comparison group nor regression discontinuity designs will not be considered responsive . . . (p. 3586)

Moreover, the DOE's What Works Clearinghouse (WWC), a government research initiative that reviews and assesses educational research, provides research summaries based on the evidence of effectiveness in meeting WWC standards of evidence based on the use of experimental methods. According to the WWC, their goal is to "use rigorous and relevant research, evaluation and statistics to improve our nation's educational system. The mission of the WWC is to be a central and trusted source of scientific evidence for what works in education."

Furthermore, the National Institute of Justice (NIJ) in the United States has also explicitly and openly declared a preference for RCTs and quasi-experimental methods in evaluation. According to the NIJ:

> The scientific validity of evaluation is measured along a continuum from strong to weak. Randomized control trials provide the strongest measure of a program's effects ... however, this "gold standard" is hard to achieve in some research situations. In these cases, we have other "quasi-experimental" methods that may provide acceptable precision in detecting and measuring the program's effects. (NIJ, 2010)

In Canada, methodological preference is articulated in the TBS *Management Accountability Framework* (1997) and in the Guide for the Development of Results Based Management and Accountability Frameworks (2010), two documents that are intended for all federal departments and that give priority to methods that focus on program relevance, success, and cost-effectiveness. Moreover, the *Policy on Evaluation* (TBS, 2009) for the Government of Canada links evaluation to results-based management, with the emphasis given to gathering evidence to demonstrate program relevance (continued need for program, alignment with government priorities, and alignment with federal roles and responsibilities) and performance (achievement of outcomes, demonstrated efficiency, and effectiveness), highlighting fiscal prudence, efficiency, and accountability (Shepherd, 2011). According to the policy, "evaluation provides Canadians, Parliamentarians, Ministers, central agencies and deputy heads an evidence-based, neutral assessment of the value for money, i.e. relevance and performance, of federal government programs" (TBS, 2009). This vision of evaluation, while not giving preference to a singular methodology as in the U.S. departments of justice and education, nonetheless articulates a vision of evaluation as an evidence-based, neutral, and objective instrument designed to provide outcome and results-based information on program performance and spending.

These American and Canadian examples, while admittedly only providing a partial picture, nonetheless illustrate clear instances of divergence between the potential ways in which evaluation could be conceptualized and utilized in our pluralized society, and the more limited ways in which evaluation is institutionalized in federal public sectors (Norris & Kushner, 2007). This discourse of evaluation (e.g., as a means of obtaining "scientific" certainty about program effects) may thus influence and delimit the potential of evaluation, both in terms of how it is conceptualized and how it is used in the evaluation of federal programs. Thus, while there continues to be considerable discussion and debate among evaluation scholars and practitioners about evaluation methodology and method use, at the level of federal policy, evaluations that are based on experimental, quasi-experimental designs and other quantitative measures continue to be given preference. In the next section, I briefly define participatory evaluation (what it is and what it looks like in practice) and then situate it within the current government climate of accountability.

## Participatory Evaluation: A Brief Description

Participatory approaches to evaluation can be defined by the engagement of participants in the evaluation process, rather than by any specific set of methods or techniques. The issue in participatory evaluation is not about which methods to use, but whose voices to include, how to include them, and determining who will speak for whom (Greene, 2000). In fact, decisions concerning method choice do not come from any a priori philosophical or methodological preference but rather from participants themselves and from the exigencies of the program and community context (Hall, 1992). As Cornwall and Jewkes (1995) have observed, participatory research can be distinguished from other approaches based on "who defines the research problems and who generates, analyses, represents, owns and acts on the information which is sought" (p. 1668). Participatory approaches to evaluation can be further differentiated from other evaluation approaches (and from each other) based on the

level and extent of stakeholder involvement, the diversity among stakeholders and the level of evaluator control over the process (Cousins & Whitmore, 1998). These three dimensions of practice, while distinct, do nonetheless overlap during the course of an evaluation, and vary depending upon the community context, the evaluators level of experience and expertise, and institutional and program influences (Cousins & Chouinard, 2012).

The rationale and motivation for adopting a participatory approach can be further distinguished based on the rationale and goals of the evaluation as well as the needs and requirements of the program and community context. Rationales used to advance collaboration include a political justification rooted in concerns for social justice and based on a moral and normative sense of obligation involving the ideals of democratic inclusion, empowerment, and emancipation (Cousins & Whitmore, 1998). A philosophic justification motivated by social constructivism and concerning the centrality of context and the inclusion of multiple participants in the production of knowledge, and a pragmatic justification that is practical in orientation, and based on the belief that increased participation will lead to results that better support program and organizational decision making (Patton, 1998) and enhance organizational learning. These three justifications, while not conceptually distinct, do impart a sense of the myriad motivations for adopting a participatory approach to evaluation, while also providing a tangible sense of the discrepancies between collaborative approaches and the way evaluation has been implemented and conceptualized in the public sector. In the next section, I discuss seven key tensions that I have identified between approaches to evaluation as it is framed within the culture of accountability in the public sector, what I have termed technocratic approaches to evaluation, and participatory and collaborative approaches to evaluation.

## Key Tensions Between Technocratic and Participatory Approaches to Evaluation

Program evaluation assumes many guises and plays many different roles, depending upon the program sponsor and whether its intended use is to render judgment, facilitate program improvement, generate knowledge, learning, accountability, ongoing monitoring, or development (Patton, 2008). Whatever the intention, evaluations can be distinguished based on what Schwandt (2009) refers to as technocratic or human-centered knowledge systems, the former as a prescriptive and instrumental approach defined by scientific knowledge claims and best practice research, and the latter by its attention to interdependence, diversity, and values. The focus of the technocratic approach is on the use of evaluation to generate a knowledge base and to provide performance planning, while the human-centered approach incorporates local community knowledge and is concerned with the needs of the local population and with self-determination. MacDonald's (1974) typology of evaluation studies (specifically bureaucratic and democratic approaches) further specifies the key distinction being advanced, between evaluations designed to maintain close control of resources and services and intended to serve accountability purposes (bureaucratic), and evaluations that value pluralism and the representation of diverse interests and voices, designed to provide service to the whole community (democratic).

These distinctions illustrate that program evaluation can play a dual role, one that legitimates and is responsive to the current system of performance measurement and accountability that is so prevalent in the public sector, and/or as a means to broaden democracy by including the voices and perspectives of the less powerful in our society and those more directly involved in the program itself. While these roles are not intended as purely distinct, as there is indeed a mixing of evaluation functions and purposes within the world of practice, in what follows I identify seven areas of tension that may potentially influence the reception of participatory or collaborative approaches to evaluation in public sector contexts. Indeed, the dichotomous characterization of the technocratic and participatory approaches that follow is intended as a heuristic to facilitate a discussion of differences between

approaches and of potential tensions that may arise. The tensions are (1) relational, (2) epistemological, (3) pedagogical, (4) contextual, (5) political, (6) methodological, and (7) organizational. Thus, while there is conceptual overlap, each tension nonetheless provides a unique perspective in which to critically examine the relationship between participatory and technocratic approaches to evaluation.

*Relational.* Participatory evaluation is a fundamentally relational and interactive approach to evaluation that is grounded in the social relations between evaluators and stakeholders. As such, evaluators and stakeholders are not considered separate but rather inextricably linked together in what Heron (1996) has referred to as "intersubjective space." The relationship itself is considered a key component of participatory research (Kelly, 2006; Trickett & Espono, 2004), as evaluators and stakeholders work together in the joint construction of evaluative knowledge (Rebien, 1996). In fact, it is the relationships that emerge among stakeholder groups, as well as the dialogues and conversations that develop, that effectively defines the parameters of participatory practice (Abma & Widdershoven, 2008). Trust is thus paramount, as evaluators and stakeholders work closely together in what are often diverse, complex, and conflictual program and community settings. As Symonette (2004) observes, "trust is the glue and fuel for cultivating viable and productive social relations" (p. 100).

A technocratic approach to evaluation, on the other hand, assumes a far more modest and detached position relationally, as accountability has redefined relations between the state and its citizens to somewhat formal relationships based on economic identities (Biesta, 2004). Relationships with the state are thus circumscribed, as schools and other governmental authorities are no longer accountable to what Biesta (2004) refers to as "real stakeholders" but to external authorities who are even further removed from the community and program context. As such, accountability does not lead to greater dialogue among relevant program stakeholders (Power, 1997), but to less, as stakeholders busily engage in endless fact checking and reporting in response to state accountability requirements. As Norris and Kushner (2007) describe:

> The social relations of the New Public Management construct sharp distinctions between the actors in public services ... mostly those subject to evaluation such as teachers have little or no say about the way in which it is done. Similarly, the intended beneficiaries of public services are largely uninvolved in setting the agenda for evaluation. (pp. 11–12)

Relationships in the field are thus structured around satisfying the plethora of accountability demands rather than in working collectively and collaboratively with colleagues and other stakeholders, a fundamentally antidemocratic admission when we consider the diversity of voices and perspectives that will remain unheard.

The role an evaluator assumes during the course of the evaluation is also very much related to the types of relationships that will be developed and tolerated during the course of an evaluation, from engaged colleague to neutral and detached judge or observer. As such, evaluator "positionality" has everything to do with how evaluation is conceptualized within a particular program setting (King & Stevahn, 2002). For example, within a participatory context, the evaluator is responsive to the requirements of the program and community setting and to the stakeholders who are involved, as well as to the rationale adopted to advance participation. The essence of the evaluator role in a participatory evaluation is thus given to creating the conditions and circumstances that will encourage and enable stakeholder participation throughout the evaluation process (Trickett & Espino, 2004).

The role of evaluators within a technocratic approach is purported to be more independent and objective, as external evaluators are hired to "perform" evaluation with little time to focus on building the kinds of relationships that some would argue are necessary in evaluation (Biott & Cook,

2000). Thus, while the credibility of the evaluation is premised on a detached and neutral role for evaluators, the irony is that many external evaluators are hired by overly involved program managers to conduct evaluations that have preordinate designs, predetermined indictors, evaluation questions, and measures of success (Gauthier et al., 2004).

*Epistemological.* Participatory knowledge claims are grounded in a social constructivist epistemology and are very much attuned to, as well as embedded within the social, historical, and political context of the program community. Knowledge is not framed in narrow technical terms and treated as a commodity that can be collected and accumulated through research means and methods. Rather, knowledge is considered to be multifaceted and multidimensional, enjoining instrumental, relational, and critical forms of knowledge construction (Habermas, 1971; Park, 1993). Instrumental knowledge, based on the natural sciences, is technical in nature and is acquired through quantitative measures that seek understanding through the principles of cause and effect, predictability, and control. It is philosophically aligned with positivism and empirical methodologies. Relational or practical knowledge is founded on the interpersonal connections among people, both on a personal level and more broadly from a sociopolitical perspective, and is motivated by our concern and understanding of others. As such, it is philosophically grounded in hermeneutics and other interpretive methodologies. Critical knowledge, what Habermas (1971) refers to as emancipatory knowledge, comes from a process of self-reflection and a critical questioning of assumptions, positions, beliefs, and values. Philosophically it is based on principles of critical theory and related to critical social science methodologies. In contradistinction to technocratic approaches to evaluation where the focus is on instrumental forms of knowledge, during a successful participatory evaluation, all three forms of knowledge (instrumental, relational, and critical) are developed simultaneously (Park, 1993).

Relational and critical forms of knowledge help to distinguish participatory approaches to evaluation from more technocratic approaches, as knowledge is considered something that is constructed collectively with stakeholders through the process of social interaction (Long, 1992) rather than through outside "experts" and facilitated by an "objective" evaluator. In the climate of accountability, knowledge claims that have not been derived through positivist/postpositivist methodologies are considered less credible (Henry, 2009; Steiner, Wroblewski, & Cook, 2009), leading to what Reagan (1996) would call a form of "epistemological ethnocentrism," where "legitimate discourse" is firmly established along rather narrow dimensions. Foucault's (1980) concept of geneology helps to illustrate the distinction between what are considered "legitimate knowledge" claims articulated through scientific discourse and "subjugated knowledge" based on local and regional community expressions. Legitimate knowledge is considered an erudite, formal type of knowledge that carries with it the authority to define the parameters of official discourse, while "subjugating" local knowledge to the periphery of epistemological acceptability (Foucault, 1980). The challenge of conducting participatory evaluations in culturally diverse communities thus remains, as culturally specific and locally meaningful knowledge that is based on local (and perhaps indigenous) epistemological constructs continues to challenge the parameters of what is considered accurate, reliable, and valid knowledge (Letiecq & Bailey, 2004). The power to define legitimate forms of knowledge is thus vested externally, residing outside of the local context, drawing attention to the politicized nature of knowledge construction in the social sciences, and its circumscribed and narrow boundaries. As Norris and Kushner (2007) so accurately observe, "knowledge control lies at the heart of all New Public Management projects" (p. 3).

*Pedagogical.* Evaluation is itself considered a pedagogical undertaking (Schwandt, 2003), as learning is implicated in the creation and use of evaluation findings and through stakeholder involvement in the process and practice of evaluation. To understand the key difference between the two approaches in terms of learning, consider Schwandt's (2003) distinction between evaluation conceptualized as a

"technical undertaking," described as a set of tools that if used correctly may improve practice and generate answers, and evaluation as a conceptual practice designed to generate dialogue and facilitate learning. Learning at the technical level is focused on the practical aspects of data collection methods and evaluation, while learning at the conceptual level moves beyond the practical and technical aspect (beyond "knowing how") to incorporate interactive, social, and dialogic learning, to what can be considered a form of "learning-in-practice" or "situated learning" (Lave & Wenger, 1991, p. 527).

Participatory evaluation is a fundamentally educational process (Oakley, 1991), considered by some as a learning system (Cousins & Earl, 1992), where stakeholders work alongside evaluators in identifying issues, carrying out research tasks, and responding to research findings and results. It is through the process of participation itself, and through the relationships that are created among evaluation participants, that learning takes place. As such, learning occurs at a practical or informational level concerning the program, the organization, the context, and the evaluation itself, as well as at a conceptual or reflective level concerning relationships to self and others (Oakley, 1991). The focus of the technocratic approach to evaluation, on the other hand, while not incompatible with evaluation for learning, emphasizes relationships of accountability and compliance with predetermined targets and standards, rather than with stakeholder involvement and capacity building through an interactive and reflective process of engagement. From the technocratic perspective, learning can be attributed to the instrumental and symbolic use of findings, to the knowledge that is created to assist program decision making and the future allocation of resources as well as to provide evidence of program quality to government stakeholders and the public. Learning from a participatory perspective, on the other hand, can be attributed to evaluation outcomes and use of findings and/or less directly through the effects of the evaluation process on participants (Cousins & Chouinard, 2012), what Patton (2008) refers to as "process use." According to Patton (2008), process use is defined as "individual changes in thinking, attitudes, and behavior, and program or organizational changes in procedures and culture that occur among those involved in evaluation as a result of the learning that occurs during the evaluation process" (p. 155). Learning from the technocratic perspective, on the other hand, is focused on tracking performance and progress against stated goals to ensure program efficiency and effectiveness, rather than as an educative tool to enhance capacity building and dialogue at the individual, group, or organization level.

*Contextual.* A participatory approach to evaluation is strongly grounded in the program and community context, as all communities are considered to have a local ecology (Kelly, 2006) that influences the dynamics of participation and situates it within a specific social, cultural, political, and institutional setting (Trickett & Espino, 2004). Participatory evaluation is thus context rich as well as context specific, as it is grounded in the concerns, interests, and problems of the program community (Burke, 1998), all of which are enmeshed in shaping the parameters of the evaluation process. Context can also be understood from a micro perspective involving the dynamics of the local community and program setting as well as from a macro perspective situating the evaluation within a broader sociopolitical context. Both micro- and macro-contextual levels influence the dynamics of the evaluation, the relationships and politics within the local setting, and the broader sociopolitical and historical influences that also include the professional norms and ideology that guide practice as well as the diversity of program needs (Cousins & Chouinard, 2012).

The technocratic approach to evaluation, on the other hand, is very much focused on macro-level applications that include predetermined metrics and universal standards, rather than on the micro-contextual level dynamics of the local program and community expressions. In fact, the centralization of evaluation and performance management functions within federal institutions has created a significant tension between meeting local departmental needs amid the exigencies of central agency accountability and fiscal demands (Shepherd, 2011). Unlike participatory evaluation where

community dynamics and the diversity of context has a profound influence on participatory methodology (Guijt & Shah, 2001), the technocratic approach to evaluation minimizes stakeholder engagement and glosses over context, thereby reducing complexity at the microlevel to a simplistic formulaic and linear process (Davies, 1999). The lack of contextual sensitivity tends to depoliticize evaluation practice by ignoring the fact that someone (and not others) was involved in the construction of these programs and policies, and that these programs and policies thus exist within contested spaces where ethnicity, gender, social class sexuality, and so on are enmeshed (Everitt, 1996; Schwandt, 2009). As such, the distinction and potential tension between micro and macro levels highlights the tension not only between the needs of the community and the needs of the sponsor (Biott & Cook, 2000; Chaytor, MacDonald, & Melvin, 2002; Wallerstein, 1999) but also between the participatory and technocratic approach to evaluation.

*Political.* Participatory evaluation can be described as a political process as much as a methodology (Gaventa, Creed, & Morrissey, 1998), with a focus on who does or does not participate, under what conditions they participate, and what form their participation takes over time. Despite the attempt to empower participants and engage local knowledge in developing, framing, and conducting the evaluation, power continues to shape the participatory context (Greene, 2000; Gregory, 2000). Power is thus conceptualized as a dynamic, relational, and productive concept (Foucault, 1979), something that is not possessed by any one individual but something that circulates and structures knowledge and its social construction at multiple levels (Foucault, 1980). Previously (Cousins & Chouinard, 2012), we have noted three key levels of power within a participatory context (relational, political, and discursive), all of which are interrelated and act at different locations within the evaluation setting. Relational power refers to the social relations within the participatory context, creating categories of people who can or cannot participate in the evaluation. Political power refers to the barriers and biases that prevent people from participating, influencing whose voices are heard and whose agendas are included. Discursive power refers to the internalized norms and values that guide our practice and that preclude some people from participating and prevent certain perspectives from being voiced and heard. In a participatory approach, all three types of power can manifest and can influence the evaluation process and eventual evaluation outcomes.

Rather than addressing differences in power, privilege, and social status between stakeholders, what Greene (2007) refers to as "engaging with difference," technocratic approaches to evaluation tend to legitimate and reinforce existing power structures (Lehtonen, 2005) by imposing strict central control. Method selection (e.g., quantitative, experimental, and quasi-experimental approaches) precludes the inclusion of multiple and diverse stakeholders, thus discouraging public discussion about which standards should be adopted (or not), and which outcomes might be most socially desirable (Biesta, 2004). By excluding alternative voices and silencing potentially diverse perspectives, technocratic approaches to evaluation thus work to depoliticize public policy decision making (Schwandt, 2009). As a result, much of the discussion remains at the level of methodology, and concerns approaches for better monitoring performance to ensure accountability, rather than about what values we want to preserve as a society, and what goals we ought to collectively and actively pursue (Stein, 2001). As Greene (2009) states, "the argument is not about method. Rather, it is about the *politics* of method—the political underbelly of the current demand for evidence-based policies and practices" (p. 158). From a Foucauldian perspective, the technologies of control that we have created to monitor our compliance to the principles of NPM, through our adherence to discourses of accountability and public sector performance measures, has created a system of self-monitoring and self-surveillance, what Foucault (1972) might have referred to as a form of disciplinary power. We thus slide unconsciously from a notion of accountability as a democratic principle denoting shared responsibility, to a predominantly technocratic conception that is focused on control, internal monitoring, and compliance, as far removed from a democratic notion as could be envisioned.

*Methodological.* Methodologies include philosophical assumptions that frame the process of social inquiry, and that help to shape the research objectives, questions, and design of the study (Carter & Little, 2007). For Schwandt (2007), methodologies are "middle grounds" between epistemologies that guide our research, and the methods that we ultimately use to gather our research findings. Regardless of rationale for collaboration, in a participatory evaluation the selection of methods is based on decisions made collaboratively by evaluators and community and program stakeholders, and informed by community and cultural needs, funder requirements, and other contextual conditions. As such, participatory approaches do not favor any one method, as decisions are made collectively around data collection and the selection of methods. As such, in a participatory evaluation, method selection and design can be described as eclectic, as methods are adapted for use based on organizational and community context, rather than on predetermined metrics and measures of success.

Despite the plethora of contextual factors to consider in selecting and designing an evaluation, from a technocratic perspective the methodology of choice is often quantitative, with experimental or quasi-experimental design given gold status and applied across multiple program and evaluation contexts. The focus on outcome measures at the expense of other important and relevant evaluation issues, what Bell (2004) refers to as the "tyranny of methods," reduces contextual complexity and diversity to narrowly observable and recordable parameters. As a result, all methods that do not meet the gold standard (e.g., objective, neutral, and evidence-based) are eschewed, regardless of program or community needs. The singular focus on performance measurement and monitoring is thus politically motivated (e.g., see Greene, 2009; Norris & Kushner, 2007; Ryan, 2007; Schwandt, 2007), as it limits stakeholder input and public dialogue, narrows program and community understanding to predetermined measures of program success, confines program quality to measurable outcomes, and circumscribes method selection. Seen from this perspective, the contrast between technocratic and participatory approaches to evaluation is stark.

*Organizational.* The organizational tension identified here is the result of managerial and institutional constraints (both structurally and philosophically) that influence the program and its subsequent evaluation. A particular challenge in participatory contexts is insufficient time or resources to develop the kinds of relationships and understanding between evaluators and program and community stakeholders necessary to conduct participatory or collaborative evaluations. As Biott and Cook (2000) note, the need to satisfy bureaucratic requirements comes at the cost of engaging local stakeholders and responding to local priorities. This tension can be further exacerbated by public sector sponsors who regard evaluation as an exercise in accountability and control, rather than as a means to learning and local community engagement. One of the key tensions is finding a balance between what can be the conflicting needs and priorities of community and program stakeholders and public sector funding agencies (Biott & Cook, 2000; Trickett & Espino, 2004). The success of a participatory evaluation thus requires a supportive organizational culture that includes the availability of resources (time and support), information and program needs that are commensurate with a participatory methodology, and a shared evaluation purpose (Cousins & Chouinard, 2012).

In contrast to a bottom–up participatory approach to evaluation that is premised on local collaboration and community input, a technocratic approach is based on a top–down hierarchical structure with central control, driven by a need for efficiency and effectiveness rather than democratic principles (Behn, 1998). As Radin (2002) notes, "the focus on hierarchical control of bureaucracy has been the approach to accountability that dominates the public administration field" (p. 14). Thus, despite the expressed need for objectivity and neutrality, the current approach to evaluation is highly prescriptive, and, rather ironically, results in organizational and managerial overinvolvement, notwithstanding the use of external evaluators. Accountability thus leads to the cooptation of management systems, turning them into monitoring systems to ensure compliance with preestablished

and externally derived organizational norms and standards (Power, 1991). The top–down approach to management reinforces hierarchical and bureaucratic relations, thus providing a challenging context for participatory evaluation to thrive (King, 1998), and while at the same time, the use of predetermined measures reduces the organizational learning potential of evaluation.

## Concluding Thoughts

In this article, I focused on the tensions between participatory approaches to evaluation with the vision of evaluation that prevails within the federal public sector, what I have referred to throughout as the technocratic approach to evaluation. The tensions identified in this article between technocratic and participatory approaches to evaluation (e.g., relational, epistemological, pedagogical, contextual, political, methodological, and organizational) provide a stark contrast between the two approaches and help illustrate some of the key challenges of introducing participatory approaches to evaluation in the prevailing climate of accountability. Given the fundamental contrasts between the two approaches (e.g., in terms of stakeholder involvement, knowledge generation, learning opportunity, contextual relevance, power and control, method preference, and program community vs. state needs), the reception and use of participatory methods in such a climate remains daunting.

Evaluation occurs within a specific context and is necessarily influenced by the economic, political, historical, and social forces that shape that context. The culture of evaluation is thus very much embedded in the culture of accountability that currently prevails in public sector institutions, policies, and programs (Stronach, Halsall, & Hustler, 2002). As such, our understanding of the reception and use of participatory approaches to evaluation must include an understanding of the practices of NPM and of the concomitant call for accountability and performance measurement standards within public sector institutions. The practice and function of evaluation is a key part of the fabric of NPM as it is expressed in public sector governance mandates, a key part of the rhetoric of public accountability. As Chelimsky (2008) explains, "evaluators derive their legitimacy, and also their role and mandate, their need for technical competence and credibility, along with their right to independence, from the political notion of accountability in government" (p. 401). Despite the apolitical guise of technocratic approaches, evaluation is not merely a technical activity bereft of political or normative values (Greene, 2002) but is deeply implicated and influenced by the political, social, historical, and economic context in which it occurs. Historically, there has always been a very strong link between evaluation and the workings of government, where evaluation is seen as a means of informing and legitimizing governmental decision making (House, 1993), thus becoming a key part of government's "legitimizing rhetoric" (Stronach et al., 2002, p. 176). As Segerholm (2002) describes, "evaluation practice is part of larger political and managerial moods in today's society and functions not only as a source of information to the public but as part of a larger web of different controlling, monitoring, and governing practices" (p. 98). Evaluation is thus a key part of what it means to manage public policy, perception, and financial expenditures under NPM, and concomitantly a key part of accountability as it is currently defined and practiced. In fact, what is becoming ever more apparent is the explicit role evaluation now plays in the state's system of governance (Ryan & Feller, 2009; Schwandt, 2009), not merely as passive bystander or reluctant companion, but as an active partner with audit and performance management in framing our expectations of government programs, policies, and public services.

In this article, I have argued for the need to create opportunities (and space) for participatory approaches to evaluation, for a place where our evaluation practices reflect the diversity and exigencies of their program contexts. As I have highlighted in section two of the article, the world of politics and the world of evaluation are never very far apart. Thus, while there remain many areas of consensus among evaluators, and discussion and debate is ongoing (see Donaldson, Christie, & Mark 2009; Julnes & Rog, 2007), disagreement is never far away, as federal public policy decisions

continue to surface methodological dissent in the field (see Donaldson & Christie, 2005). Given the culture of accountability that prevails in much of the Western world (including the predominance of NPM throughout the public sector), the *rough ground of practice* is not an *even ground*, not a place where evaluation methodologies are provided an equal footing.

While much of the discussion revolves around the technical merits of experimental, quasi-experimental, or quantitative methods versus other inquiry approaches, the fundamental debate has less to do with technical quality than with what Schwandt (2007) terms "epistemological politics" (the privileging of some knowledge and some knowers over others). In a society defined by deep diversity (in terms of ethnicity, class, gender, etc.), the adoption and use of standardized "scientific" measures must also be regarded as a form of "epistemological ethnocentrism" (Reagan, 1996), framing and legitimizing the very parameters, the assumptions and biases, within which evaluation takes place. As Danziger (quoted in Kushner, 2000) explains, "forms of investigation have to be understood as forms of social organization-they do not happen outside of our institutions, our social relations, our politics, and economics" (p. 16). Danziger (1990) thus draws a clear distinction between the methodologies that we use and the actual discourses that circulate about the methods and approaches that we adopt for use in the field. Questions of epistemology (what knowledge is, how we create it, what we do with it, how it circulates, and how it is transformed) thus become paramount if we are to move beyond debates about technical merit and method selection, to more specific interrogations of the political and economic structures that shape our methodologies and our experiences as social inquirers and evaluators in the field. Thus, while NPM and accountability are a core part of the political fabric of government program policy and decision making, evaluators can nonetheless create opportunities for dialogue, engagement, and education about what evaluation is, and the role that it can play in fostering a more democratic and culturally responsive program context.

## Declaration of Conflicting Interests

## Funding

## References

Abma, T. A., & Widdershoven, G. A. M. (2008). Sharing stories: Narrative and dialogue in responsive nursing evaluation. *Evaluation & the Health Professions*, *28*, 90–109.

Behn, R. D. (1998). The new public management paradigm and the search for democratic accountability. *International Public Management Journal*, *1*, 131–164.

Behn, R. D. (2001). *Rethinking democratic accountability*. Washington, DC: Brookings Institute Press.

Bell, S. (2004). Methods and mindsets: Towards an understanding of the tyranny of methodology. *Public Administration & Development*, *14*, 323–338.

Biesta, G. J. J. (2004). Educational accountability and the ethical demand: Can the democratic potential of accountability be regained? *Educational Theory*, *54*, 233–250.

Biott, C., & Cook, T. (2000). Local evaluation in a national early years excellence centres pilot programme: Integrating performance management and participatory evaluation. *Evaluation*, *6*, 399–413.

Blalock, A. B. (1999). Evaluation research and the performance management movement: From estrangement to useful integration. *Evaluation*, *5*, 117–149.

Burke, B. (1998). Evaluating for change: Reflections on participatory methodology. In E. Whitmore (Ed.), *Understanding and practicing participatory evaluation: New directions in evaluation, No. 80* (pp. 43–56). San Francisco, CA: Jossey-Bass.

Campbell, R. F., Fleming, T., Newell, L., & Bennion, J. W. (1987). *A history of thought and practice in educational administration*. New York, NY: Teachers College Press.

Carter, S. M., & Little, M. (2007). Justifying knowledge, justifying method, taking action: Epistemologies, methodologies, and methods in qualitative research. *Qualitative Health Research*, *17*, 1316–1328.

Charlton, B. G. (2002). Audit, accountability, quality and all that: The growth of managerial technologies in UK universities. In S. Prickett & P. Erskine-Hill (Eds.), *Education! education! education!—Managerial ethics and the law of unintended consequences* (pp. 13–28). Thorverton, UK: Imprint Academic.

Chaytor, K., MacDonald, G., & Melvin, D. (2002). Preparing nonprofits for new accountability demands. *The Canadian Journal of Program Evaluation*, *17*, 93–112.

Chelimsky, E. (2008). A clash of cultures: Improving the "fit" between evaluative independence and the political requirements of a democratic society. *American Journal of Evaluation*, *29*, 400–415.

Cornwall, A., & Jewkes, R. (1995). What is participatory research? *Social Science & Medicine*, *41*, 1667–1676.

Cousins, J. B., & Chouinard, J. A. (2012). *Participatory evaluation up close: A review and integration of the research base*. Greenwich, CT: Information Age Publishing Inc.

Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis*, *14*, 397–418.

Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. In E. Whitmore (Ed.), *Understanding and practicing participatory evaluation: New directions in evaluation, no. 80* (pp. 3–23). San Francisco, CA: Jossey-Bass.

Cronbach, L. J., Ambron, S. W., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., Walker, D. F., & Weiner, S. S. (1980). *Toward a reform of program evaluation*. San Francisco, CA: Jossey-Bass.

Davies, I. C. (1999). Evaluation and performance management in government. *Evaluation*, *5*, 150–159.

Donaldson, S. I., & Christie, A. A. (2005). Determining causality in program evaluation and applied research: Should experimental evidence be the gold standard? *Journal of Multidisciplinary Evaluation*, *3*, 60–77.

Donaldson, S. I., Christie, C. A., & Mark, M. M. (Eds.). (2009). *What counts as credible evidence in applied research and evaluation practice?* Thousand Oaks, CA: Sage.

Eisner, E. W. (1985). *The educational imagination: On the design and evaluation of school programs* (2nd ed.). New York, NY: Macmillan.

Everitt, A. (1996). Developing critical evaluation. *Evaluation*, *2*, 173–188.

Fitzpatrick, J., Christies, C., & Mark, M. M. (2009). *Evaluation in action: Interviews with expert evaluators*. Thousand Oaks, CA: Sage.

Foucault, M. (1972). *The archeology of knowledge and the discourse on language*. New York, NY: Pantheon Books.

Foucault, M. (1979). What is an author? *Screen*, *20*, 13–33.

Foucault, M. (1980). *The history of sexuality*, *Vol. 1*. New York, NY: Vintage.

Franklin, U. (1990). *The real world of technology*. Toronto, Canada: CBC Enterprises.

Gauthier, B., Barrington, G., Bozzo, S. L., Chaytor, K., Cullen, J., Lahey, R. . . . Roy, S. (2004). The lay of the land: Evaluation practice in Canada today. *Canadian Journal of Program Evaluation*, *19*, 143–178.

Gaventa, J., Creed, V., & Morrissey, J. (1998). Scaling up: Participatory monitoring and evaluation of a federal empowerment program. In E. Whitmore (Ed.), *Understanding and practicing participatory evaluation: New directions in evaluation, no. 80* (pp. 81–94). San Francisco, CA: Jossey-Bass.

Greene, J. C. (1999). The inequality of performance measurements. *Evaluation*, *5*, 160–172.

Greene, J. C. (2000). Challenges in practicing deliberative democratic evaluation. In K. E. Ryan & L. DeStefano (Eds.), *Evaluation as a democratic process: Promoting inclusion, dialogue, and deliberation, no. 85* (pp. 27–38). San Francisco, CA: Jossey-Bass.

Greene, J. C. (2005). Evaluators as stewards of the public good. In S. Hood, R. Hopson, & H. Frierson (Eds.), *The role of culture and cultural context: A mandate for inclusion, the discovery of truth, and understanding in evaluative theory and practice* (pp. 7–20). Greenwich, CT: Information Age Publishing.

Greene, J. C. (2007). *Mixed methods in social inquiry*. San Francisco, CA: Jossey-Bass.

Greene, J. C. (2009). Evidence as ''proof'' and evidence as ''inkling''. In S. I. Donaldson, C. A. Christie, & M. M. Mark (Eds.), *What counts as credible evidence in applied research and evaluation practice?* (pp. 153–167) Thousand Oaks, CA: Sage.

Gregory, A. (2000). Problematizing participation: A critical review of approaches to participation in evaluation theory. *Evaluation*, *6*, 179–199.

Guijt, I., & Shah, M. K. (2001). Waking up to power, conflict and process. In I. Guijt & M. K. Shah (Eds.), *The myth of community: Gender issues in participatory development* (pp. 1–23). London, UK: ITDG Publishing.

Habermas, J. (1971). *Knowledge and human interests*. Boston, MA: Beacon Press.

Hall, S. (1992). The west and the rest: Discourse and power. In S. Hall & B. Gieben (Eds.), *Formations of modernity* (pp. 275–331). Cambridge, UK: Polity Press.

Henry, G. T. (2009). When getting it right matters: The case for high-quality policy and program impact evaluations. In S. I. Donaldson, C. A. Christie, & M. M. Mark (Eds.), *What counts as credible evidence in applied research and evaluation practice?* (pp. 32–50). Thousand Oaks, CA: Sage.

Heron, J. (1996). *Co-operative inquiry: Research into the human condition*. London, UK: Sage.

House, E. R. (1993). *Professional evaluation: Social impact and political consequences*. Thousand Oaks, CA: Sage.

Julnes, G., & Rog, D. J. (Eds.). (2007). *Current federal policy and controversy over methodology in evaluation: New directions for evaluation, no. 113*. San Francisco, CA: Jossey-Bass.

Kelly, J. G. (2006). *Becoming ecological: An expedition into community psychology*. New York, NY: Oxford University Press.

King, J. A. (1998). Making sense of participatory evaluation practice. In E. Whitmore (Ed.), *Understanding and practicing participatory evaluation: New directions in evaluation, no. 80* (pp. 57–68). San Francisco, CA: Jossey-Bass.

King, J. A., & Stevahn, L. (2002). Three frameworks for considering evaluator role. In K. E. Ryan & T. A. Schwandt (Eds.), *Exploring evaluator role and identity* (pp. 1–16). Greenwich, CT: Information Age Publishing Inc.

Kushner, S. (2000). *Personalizing evaluation*. Thousand Oaks, CA: Sage.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York, NY: Cambridge University Press.

Lehtonen, M. (2005). OECD environmental performance review programme: Accountability (f)or learning. *Evaluation*, *11*, 169–188.

Letiecq, B. L., & Bailey, S. J. (2004). Evaluating from the outside: Conducting cross-cultural evaluation research on an American Indian reservation. *Evaluation Review*, *28*, 342–357.

Long, N. (1992). Introduction. In N. Long & A. Long (Eds.), *Battlefields of knowledge: The interlocking of theory and practice in social research and development* (pp. 3–43). London, UK: Routledge.

MacDonald, B. (1974). Evaluation and the control of education. B. MacDonald & R. Walker (Eds.), *Innovation, evaluation, research and the problem of control (SAFARI)*, (pp. 37–48). , Norwich, UK: Centre for Applied Research in Education, University of East Anglia.

Mark, M. M. (2003). Toward a integrative view of the theory and practice of program and policy evaluation (Chapter 11). In S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problems: Visions for the new millennium* (pp. 183–204). Mahwah, NJ: Lawrence Erlbaum Associates.

Mayne, J., & Rist, R. C. (2006). Studies are not enough: The necessary transformation of evaluation. *The Canadian Journal of Program Evaluation*, *21*, 93–120.

McDavid, J. C., & Huse, I. (2006). Will evaluation prosper in the future? *The Canadian Journal of Program Evaluation*, *21*, 47–72.

Meyer, J. W., & Rowan, B. (1983). Institutionalized organizations: Formal structure as myth and ceremony (Chapter 3). In W. Foster (Ed.), *Loose coupling revisited: A critical view of Weick's contribution to educational administration* (pp. 75–94). Victoria, Australia: Deakin University Press.

Norris, N., & Kushner, S. (2007). The new public management and evaluation. In S. Kushner & N. Norris (Eds.), *Dilemmas of engagement: Evaluation and the new public management. Advances in Program Evaluation* (Vol. *10*, pp. 1–16). Oxford, UK: Elsevier.

Oakley, P. (1991). *Projects with people: The practice of participation in rural development*. Geneva, Switzerland: International Labour Organization.

Office of Management and Budget. (2012). *Government Performance and Results Act of 2010*. Washington, DC: U. S. Goverment.

O'Neill, O. (2002). A question of trust, *BBC Reith Lectures*. Retrieved October 2, 2011, from http://www.bbc.co.uk/radio4/reith2002

Park, P. (1993). What is participatory research? A theoretical and methodological perspective. In P. Park, M. Brydon-Miller, B. L. Hall, & T. Jackson (Eds.), *Voices of change: Participatory research in the U.S. and Canada* (pp. 1–20). Westport, CT: Bergin and Garvey.

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.

Perrin, B. (2007). Towards a new view of accountability. In M.-J. Bemelmans-Videc, J. Lonsdale, & B. Perrin (Eds.), *Making accountability work: Dilemmas for evaluation and for audit* (pp. 41–59). New Brunswick, NJ: Transaction Publishers.

Power, M. (1997). *The audit society: Rituals of verification*. New York, NY: Oxford University Press.

Radin, B. A. (2002). *The accountable juggler: The art of leadership in a federal agency*. Washington, DC: CQ Press.

Reagan, T. (1996). *Non-western educational traditions: Alternative approaches to educational thought and practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

Rebien, C. C. (1996). Participatory evaluation of development assistance: Dealing with power and facilitative learning. *Evaluation*, *2*, 151–171.

Ryan, K. E. (2007). Changing contexts and relationships in educational evaluation. In S. Kushner & N. Norris (Eds.), *Dilemmas of engagement: Evaluation and the new public management. Advances in program evaluation* (Vol. *10*, pp. 103–115). Oxford, UK: Elsevier.

Ryan, K. E., & Feller, I. (2009). Education, accountability, and performance measurement in national education systems: Trends, methods, and issues. In K. E. Ryan & J. B. Cousins, (Eds.), *The Sage handbook of educational evaluation* (pp. 171–189). Thousand Oaks, CA: Sage.

Schwandt, T. A. (2003). 'Back to the rough ground!' Beyond theory to practice in evaluation. *Evaluation*, *9*, 353–364.

Schwandt, T. A. (2007). *The Sage dictionary of qualitative inquiry* (3rd ed.). Thousand Oaks, CA: Sage.

Schwandt, T. A. (2009). Globalizing influences in the western evaluation imaginary. In K. E. Ryan & J. B. Cousins (Eds.), *The Sage handbook of educational evaluation* (pp. 19–36). Thousand Oaks, CA: Sage.

Segerholm, C. (2002). Evaluation as responsibility, conscience, and conviction. In K. E. Ryan & T. A. Schwandt (Eds.), *Exploring evaluator role and identity* (pp. 87–102). Greenwich, CT: Information Age Publishing Inc.

Shepherd, R. P. (2011). In search of a balanced Canadian federal evaluation function: Getting to relevance. *Canadian Journal of Program Evaluation*, *26*, 1–45.

Smith, N. L. (2008). Fundamental issues in evaluation. In N. L. Smith & P. R. Brandon (Eds.), *Fundamental issues in evaluation* (pp. 1–23). New York, NY: Guilford Press.

Stake, R. (2001). A problematic heading. *American Journal of Evaluation*, *22*, 349–354.

Stein, J. G. (2001). *The cult of efficiency*. Toronto, Canada: House of Anansi Press Limited.

Steiner, P. M., Wroblewski, A., & Cook, T. D. (2009). Randomized experiments and quasi-experimental designs in educational research. In K. E. Ryan & J. B. Cousins (Eds.), *The Sage handbook of educational evaluation* (pp. 75–95). Thousand Oaks, CA: Sage.

Stronach, I., Halsall, R., & Hustler, D. (2002). Future imperfect: Evaluation in dystopian times. In K. E. Ryan & T. A. Schwandt (Eds.), *Exploring evaluator role and identity* (pp. 167–192). Greenwich, CT: Information Age Publishing Inc.

Symonette, H. (2004). Walking pathways toward becoming a culturally competent evaluator: Boundaries, borderlands, and border crossings. *New Directions for Evaluation*, *102*, 95–109.

Treasury Board Secretariat (TBS). (1997). *Management accountability framework*. Ottawa, Canada: Treasury Board Secretariat, Government of Canada.

Treasury Board Secretariat. (2006). *Federal accountability act*. Ottawa, Canada: Treasury Board Secretariat, Government of Canada.

Treasury Board Secretariat. (2009). *Policy on evaluation*. Ottawa, Canada: Treasury Board Secretariat, Government of Canada.

Trickett, E. J., & Ryerson-Espino, S. L. (2004). Collaboration and social inquiry: Multiple meanings of a construct and its role in creating useful and valid knowledge. *American Journal of Community Psychology*, *34*, 1–69.

U.S. Department of Education. (2005, January 25). Scientifically based evaluation methods. *Federal Register*, *70*. RIN 1890-ZA00.

U.S. Government Accountability Office. (2012). *Designing Evaluations (2012 Revisions)*. GAO-12-208G.

U.S. National Institute of Justice. (2010). *NIJ tests and evaluates programs, practices and equipment*. Retrieved July 10, 2012, from www.nij.gov/nij/about/testing-evaluation.htm

Wallerstein, N. (1999). Power between evaluator and community: Research relationships within New Mexico's healthier communities. *Social Science & Medicine*, *49*, 39–53.

What Works Clearinghouse. (2002). Retrieved July 10, 2012, http://ies.ed.gov/ncee/wwc.