Evaluation models and evaluation use

Damien Contandriopoulos and
University of Montreal, Canada

Astrid Brousselle
University of Sherbrooke, Canada

Abstract

The use of evaluation results is at the core of evaluation theory and practice. Major debates in the field have emphasized the importance of both the evaluator's role and the evaluation process itself in fostering evaluation use. A recent systematic review of interventions aimed at influencing policy-making or organizational behavior through knowledge exchange offers a new perspective on evaluation use. We propose here a framework for better understanding the embedded relations between evaluation context, choice of an evaluation model and use of results. The article argues that the evaluation context presents conditions that affect both the appropriateness of the evaluation model implemented and the use of results.

Keywords

evaluation use; evaluation model; evaluation theory; systematic review

Introduction

How to foster results use? This question consistently occupies center stage in evaluation debates, which is hardly surprising, since use is at the very core of every evaluation endeavor. Despite the abundance of literature on this topic, we believe there are new insights to be gained from lessons developed in other disciplinary fields.

This article presents findings from a recently published large-scale systematic review on interventions aimed at influencing policy-making or organizational behavior through knowledge exchange (Contandriopoulos et al., 2010). We used these findings to develop a framework for understanding the use of evaluation results.

First, in order to explain the origin of the analytical framework and present its main dimensions, we summarize some results of the systematic review. Then we apply the framework to four well-known evaluation models (utilization-focused evaluation, empowerment evaluation, realistic evaluation and democratic evaluation) to illustrate how it can shed new light on the understanding of results use. In conclusion, we discuss important

consequences of this new perspective on the use of evaluation results, on the choice of an evaluation model and on possible avenues for fostering results use.

Part I: Building a conceptual model of use

Data: A systematic literature review

The systematic review we used integrated results from various disciplinary perspectives on collective-level knowledge exchange. We use the term ‘collective-level’ here to describe interventions occurring at the organizational level or in policy-making arenas, as distinct from interventions targeting modification of individual behavior. This review focused on two relatively autonomous bodies of literature on knowledge exchange. The first is derived mostly from three sources: debates about the role of social sciences in society; studies of the use of evaluation results; and, to a lesser extent, rationalist management perspectives. The second body of literature, from the field of political sciences, comprises theoretical and empirical work on lobbying and on the functioning of policy networks.

Among the review’s main findings was the identification of three core dimensions of knowledge-exchange contexts: cost-sharing, polarization and social structuring. The detailed results and in-depth description of the methodology have been published elsewhere (Contandriopoulos et al., 2010). The present article is based on the first two of these three core dimensions.

Without entering into a full presentation of the systematic review, a few words on the methodology may be useful. Very schematically, for this review we developed a non-keyword based approach, which we termed ‘double-sided systematic snowball’. In the first (prospective) snowball step, we identified heuristically through team discussion seven traditions that contributed significantly to the understanding of knowledge exchange and use processes. We then identified articles perceived as representative of each tradition, generating a short list of seminal papers ($n = 34$). We used the *ISI Web of Science Citation Index* to identify all documents that cited them ($n = 4102$). Those documents were triaged according to their contribution to the understanding of the phenomenon under review, and 102 relevant documents were selected for in-depth analysis. In a second (retrospective) snowball step, we pooled the bibliographies of those selected documents ($n = 5622$) and applied algorithms to produce a set of documents consisting of articles cited five times or more and books cited seven times or more. Our objective was to identify what Greenhalgh et al. (2004) called ‘landmark papers’. In all, we analyzed 205 documents to build a new integrated model of knowledge exchange and use. Details on the methodology used can be found in (Contandriopoulos et al., 2010). Our analysis was based on methods inspired by Forbes and Griffiths’ (2002: 144) ‘analytical or theoretical synthesis’, which is close to Pawson et al.’s (2005) ‘realist review’ approach.

Collective knowledge exchange, creation and use

As stated above, the review was focused on knowledge exchange occurring at the organizational or policy-making levels. Such collective systems are characterized by high levels of interdependency and interconnectedness among participants (Jordan and Maloney,

1997; March, 1988; March and Olsen, 1976; Pressman and Wildavsky, 1973; Rhodes, 1990). In these systems, there is no single source of information and nor any passive receptors. All participants receive information from various sources, make sense of it, modify it and produce new information aimed at others. In such contexts, knowledge use depends on sense-making (Nonaka, 1994; Weick, 1995) and coalition-building (Heaney, 2006; Lemieux, 1998; Salisbury et al., 1987), as well as on persuasion and rhetoric (Majone, 1989; Milbrath, 1960; Perelman and Olbrechts-Tyteca, 1969; Russell et al., 2008; Van de Ven and Schomaker, 2002). In our view, this definition of collective-level knowledge exchange closely fits the reality of most settings of program evaluation results use.

The nature of collective systems suggests that scientific evidence seldom, if ever, directly solves organizational or policy-level problems (Dobrow et al., 2004; Elliott and Popay, 2000; Lomas, 1990; Sabatier, 1978). To be relevant, usable and meaningful, evidence needs to be embedded in what political science calls 'policy options' and what we generically describe here as action proposals. Action proposals are assertions that employ rhetoric to embed information into arguments to support a causal link between a given course of action and anticipated consequences (Bardach, 1984: 136; Brunsson, 1982; Haas, 1992; Knott and Wildavsky, 1980: 547; Majone, 1989: 6; Smith, 1984; Van de Ven and Schomaker, 2002). We thus suggest a definition of *collective-level knowledge use as the process by which users incorporate specific information into action proposals to influence others' thought, practices and collective action rules*. Since the practical capacity of any given user to influence the collective system within which he or she intervenes is extremely contingent on context-specific factors, this definition dissociates knowledge use from actual practices or outcomes (Henry and Mark, 2003; Weiss, 1978: 35).

Two core dimensions of the context of use

Ideology and polarization—Converging theoretical and empirical data on knowledge use suggest that, when a user's understanding of the implications of a given piece of information runs contrary to his or her opinions or preferences, this information will be ignored, contradicted, or, at the very least, subjected to strong skepticism and low use (Albaek, 1995; Booth, 1990; Bowen and Zwi, 2005; Bryant, 2003; Caplan et al., 1975; Elliott and Popay, 2000; Florio and Demartini, 1993; Freeman, 2007; Gabbay et al., 2003; Knott and Wildavsky, 1980; Lomas, 1990; Shulha and Cousins, 1997; Weiss, 1983; Weiss and Bucuvalas, 1980a, 1980b; Whiteman, 1985). It should also be stressed that real-life contexts of use are generally characterized by the interaction of several users, who should not be presumed to have similar perceptions about any given piece of information. This introduces the notion of issue polarization (Bourgeois and Nizet, 1993; Weiss, 1977c). Contexts are said to be characterized by low issue polarization when potential users share similar opinions and preferences regarding: 1) the problematization of the issue (consensus on the perception that a given situation is a problem and not the normal or desirable state of affairs); 2) the prioritization and salience of the issue (as compared with other potential issues); and 3) the criteria against which potential solutions should be assessed. Conversely, as the level of consensus on those aspects diminishes, issue polarization grows.

There is broad consensus in the literature reviewed that issue polarization is a core feature of the context of use. Discussions within knowledge exchange systems can take the form of technical debates based on rational processes only if the context is characterized by low issue polarization. Conversely, as the level of consensus among participants drops, polarization increases and the potential for resolving differences through rational arguments diminishes as debates tend toward a political form wherein the goal is not so much to convince the other as to impose one's opinion.

However, the literature reviewed is sharply divided on how knowledge exchange interventions should adapt to variations in issue polarization. There is a clearly perceptible bias in the evaluation literature in favor of instrumental use as opposed to symbolic use (Knorr, 1977; Whiteman, 1985). Since the more a context is polarized, the less it is seen as propitious for instrumental use through either a problem-driven or knowledge-driven model (Weiss, 1979), much of the evaluation-based literature tends to suggest that as polarization grows the potential for use diminishes (Knott and Wildavsky, 1980; Lynn, 1978; Weiss and Bucuvalas, 1980a). On the other hand, the literature rooted in political science takes a high level of polarization as a given. In this perspective, the way in which the system is polarized is the key to understanding the nature of ongoing coalitions (Carpenter et al., 2004; Hall and Deardorff, 2006; Hecl, 1978; Heinz et al., 1993; Jordan and Maloney, 1997; Kickert et al., 1999; Klijn, 1996; Rhodes, 1990), the level of involvement of the various actors (Ainsworth, 1993; Baumgartner and Leech, 1996; Coglianese et al., 2004; Epstein and Ohalloran, 1995; Larocca, 2004; Sloof and Van Winden, 2000) and the content of the information exchanged (Austen-Smith and Wright, 1992; Burstein and Hirsh, 2007; Phillips and Phillips, 1984). This tallies with observations made in the field of evaluation about the influence of ideological proximity on knowledge exchange processes (Caplan, 1979; Havelock, 1969; Weiss, 1977b, 1983; Weiss and Bucuvalas, 1980a, 1980b). Such a view conceives of information as a commodity in political struggles, with both a price and a value (Hall and Deardorff, 2006), and thus posits that it should be offered to allies and used strategically against opponents. This observation is supported by strong empirical data about the way lobbyists and organizations exchange information, the structure of communication networks between them and how the convergence of preferences and interests influences the communication process (Heaney, 2006; Heinz et al., 1993; Lowery et al., 2005; Smith, 1995; Wright, 1990).

Cost-sharing equilibrium in knowledge exchange systems

The idea that knowledge has both a cost and a value was used by Bardach to suggest that knowledge will reach those 'for whom the utility of having it exceeds the disutility of obtaining it' (1984: 126). Although this statement is highly rationalist, there is a widely shared broader assumption in the literature that actors will invest energy and resources in knowledge exchange processes to the extent that they perceive this investment to be profitable (Amara et al., 2004; Austen-Smith and Wright, 1992; Black, 2001; Campbell et al., 2009; Carpenter et al., 2003; Coglianese et al., 2004; Harries et al., 1999; Jacobson et al., 2005; Knott and Wildavsky, 1980; Landry et al., 2001; Landry et al., 2003; Larocca, 2004; Olson, 1965). If we accept the principle that knowledge exchange activities imply costs (e.g. time, money, attention), it follows that some within the knowledge exchange system will

have to incur those costs. The challenge, then, is to understand the cost-sharing equilibrium between participants. Likewise, according to the political science literature, a concentration of either benefits or losses (Baumgartner and Leech, 1996; Olson, 1965) or a high level of polarization will turn many actors into de facto lobbyists (intermediaries) investing to advocate in favor of specific action proposals.

We based the framework used here on a commonly accepted typology of participants that distinguishes users, at one end, from producers or intermediaries at the other. Users are those participants in the system who hold institutionally sanctioned positions that allow them to intervene in the practices, rules and functioning of organizational, political or social systems. Producers are participants who contribute to legitimate knowledge production institutions without having the capacity to put the knowledge developed to use (Arendt, 1972; Caplan, 1979; Rich, 1979). In the context of the present article, producers can be thought of mostly as evaluators. It is also crucial to remember that evaluators do not diffuse their results in a vacuum. Especially in polarized contexts, all kinds of stakeholders, lobbies and participants will want to have their say and will contribute to the information flow. Those third parties are variously called conveyors (Havelock, 1969), brokers (Weiss, 1977a), intermediaries (Huberman, 1994) or lobbyists (Milbrath, 1960, 1963).

We therefore suggest that any operational system of knowledge exchange and use will be characterized by a given cost-sharing equilibrium between users, on one hand, and producers or intermediaries, on the other. For example, one well-known user-centered equilibrium is consultancy, in which users assume most of the costs, as they hire and pay the consultant and usually devote attention to the results. At the other end of the continuum, another well-known equilibrium is the push model, in which the results of research funded by a third party are more or less actively brought to the attention of potential users. In this model, it is either the producers or interested intermediaries who will devote time and resources to catch users' attention.

As stated in the introduction, knowledge exchange is at the core of any evaluation endeavor. However, as the focus of our literature review suggested, knowledge exchange is a broader phenomenon. Any effort to translate evaluation results into a modification of practices or a decision must take into account the fact that use processes have to contend with competition between various sources of advice. Collective knowledge exchange systems are networks of numerous individuals with often divergent opinions, preferences and interest. At any point in time, various kinds and sources of information compete for scarce resources – namely, users' attention (Lindblom, 1959: 84; Simon, 1971: 40–1). Furthermore, the literature reviewed provides convincing arguments to the effect that users' responsibility is to balance various kinds of information (see for example the typologies of information from Hall and Deardorff, 2006; Peterson, 1995; Phillips and Phillips, 1984; Sabatier, 1978).

When conceptualizing the processes by which the influence of evaluation results is positioned in relation to that of other sources of opinion and advice, it should be noted that our review found no credible empirical data showing any positive link between level of use and information's internal validity or the conformity of its production process with scientific procedures (Albaek, 1995; Booth, 1990; Bowen and Zwi, 2005; Bryant, 2003; Caplan et al.,

1975; Elliott and Popay, 2000; Florio and Demartini, 1993; Freeman, 2007; Shulha and Cousins, 1997; Weiss, 1983; Weiss and Bucuvalas, 1980a; Whiteman, 1985). There is general agreement that the use of knowledge is influenced by its relevance (timeliness, salience and actionability), credibility and accessibility (Amara et al., 2004; Beyer and Trice, 1982; Campbell et al., 2009; Caplan et al., 1975; Cousins and Leithwood, 1986; Lavis et al., 2003; Lindblom and Cohen, 1979; McNie, 2007; Mitton et al., 2007; Sabatier, 1978; Weiss and Bucuvalas, 1980a). However, relevance and credibility are mostly perceptual dimensions that are very much influenced by users' pre-existing opinions and preferences (Gabbay et al., 2003; Greenberg and Mandell, 1991; Huberman, 1987; McNie, 2007; Milbrath, 1960; Weiss and Bucuvalas, 1980b).

Taken together, the polarization and cost-sharing dimensions presented above can be used to draw a two-axis space. In the following sections, we summarize some well-known evaluation models and discuss the positions they are likely to occupy in this space.

Part II: Contrasting evaluation models

We believe the two-dimensional framework based on polarization and cost-sharing can offer interesting insights into evaluation use and particularly into the embedded relations between the evaluation context, the choice of an evaluation model, the role of the evaluator and evaluation results use, which we discuss in Part III. In this section, to lay the groundwork for that discussion and to provide hands-on examples of how the framework can be applied to the analysis of evaluation result use, we apply it to the comparative analysis of some well-known evaluation models.

In 1991, Shadish and colleagues suggested that all evaluation models could be defined according to five core components that any evaluation model needs to address, albeit differently in each case (Shadish et al., 1991). Predictably, use is one of them, with the others being social programming, knowledge, values and practice. The aim of this classification was to better understand the core principles of the various evaluation theories in order to contrast their defining principles and characteristics. Since this seminal publication, other classifications have been suggested. For example, Alkin and Christie proposed a tree-shaped graphic in which three of the four categories proposed by Shadish et al. (1991) are used to position evaluation models (Alkin and Christie, 2004). Although less exhaustive than the approach of Shadish et al. (1991), as it does not consider the dimensions of social programming and practice, this presentation is interesting in that it classifies models according to their dominant component. More recently, Stufflebeam and Shinkfield (2007) proposed another classification for interpreting and classifying evaluation models. They suggested four categories: pseudoevaluations, improvement/accountability approaches, social agenda/advocacy approaches and eclectic approaches. For a comprehensive and contrasted test sample to illustrate the application of our framework, we chose models that appear on each of the different branches of Alkin's and Christie's tree and that relate, at the same time, to each category in Stufflebeam's and Shinkfield's model. We selected models that, according to Alkin's and Christie's classification, prioritize use: utilization-focused evaluation and empowerment evaluation. We chose these two approaches because they are soundly discussed and contrasted in Stufflebeam's and Shinkfield's book: utilization-

focused evaluation is classified as an eclectic approach, while empowerment evaluation is classified as a pseudo-evaluation theory. Similarly, we chose realistic evaluation because it was positioned as a methods-oriented approach according to Alkin's and Christie's classification and because it belongs to the improvement/accountability category in Stufflebeam's and Shinkfield's model. Finally, we chose democratic evaluation, which is placed on the valuing branch (Alkin and Christie, 2004) and in the social agenda/advocacy approaches (Stufflebeam and Shinkfield, 2007). While this sample clearly does not do full justice to the diversity and complexity of the evaluation field, it nevertheless offers contrasted approaches to evaluation (see table 1).

Positioning some well-known evaluation models in our framework

To illustrate how our two-dimensional typology can be used, in the following paragraphs we propose the relative positioning of some well-known evaluation models. However, we preface our discussion with two disclaimers. First, models are intellectually slippery beasts; whenever one of us felt we had gained a secure grasp on where a model should be positioned, the other would present a different interpretation that would re-open the question. While we managed to reach consensus, we do not pretend to offer definitive answers. As we explain in the third section of this article, our ambition is not to position these evaluation models definitively, but rather to offer trans-theoretical and innovative insights into the relationships between use, models and contexts. Second, although we have positioned the models in two-dimensional matrices, we must stress that the number of squares they occupy has nothing to do with their respective usefulness or the breadth of their relevance.

Utilization-focused evaluation (Patton)—Use is central to Patton's utilization-focused evaluation (UFE), which 'begins with the premise that evaluations should be judged by their utility and actual use; therefore, evaluators should facilitate the evaluation process and design any evaluation with careful consideration of how everything that is done, *from beginning to end*, will affect use' (Patton, 1997: 20). The goal of UFE is intended use (of the evaluation results) by intended users (Patton, 2005). All steps of the evaluation are thought out with the aim of facilitating use by intended users, commonly called stakeholders. Patton defines stakeholders as 'people who have a stake - a vested interest - in evaluation findings' (Patton, 1997: 41). The first step in UFE consists in identifying the intended users of the evaluation results and working closely with them to draw the evaluation parameters. By definition, this implies that users have committed, from the outset, to invest money and attention in the evaluation process. This position places UFE on the user pole of the cost-sharing equilibrium axis.

Positioning UFE in terms of polarization is not as straightforward. One interpretation considers that UFE is aimed at building consensus among stakeholders and fosters only instrumental use (as opposed to other types of use (Beyer and Trice, 1982)), in which case it would be hard to conceive of how UFE could be implemented outside low-polarized contexts. Another interpretation, based on a more restrictive view of users, posits that if the users are a subset of all stakeholders, and if those users could appropriate evaluation findings to strategically defend their own preferences and interests, then UFE can be implemented in

any context regardless of the level of polarization. In Patton's description of UFE, the aim of fostering instrumental use (Patton, 2005) clearly dominates over strategic (Greenberg and Mandell, 1991; Whiteman, 1985) or symbolic (Beyer and Trice, 1982; Knorr, 1977; Weiss and Bucuvalas, 1980a) use. Nevertheless, one could interpret this stance as aiming to increase instrumental use of the evaluation in a specific group of primary users in order to, in a second step, play a strategic role in a broader polarized arena. Our own understanding of UFE tallies with this second view and is illustrated in Figure 1.

Realistic evaluation (Pawson and Tilley)—Realistic evaluation is a variant of theory-driven evaluation. It considers programs as theories and suggests that the evaluator's role is to understand what works, for whom, and in which circumstances (Pawson and Tilley, 1997, 2005). It is a method-driven evaluation, in Alkin's and Christie's (2004) classification, and belongs to the improvement/accountability approaches in the typology of Stufflebeam and Shinkfield (2007). This model is not centered on the needs of specific users. It aims rather at providing a deep understanding of program theory centered on the interaction between the program and its specific contextual conditions of implementation. It may be funded either by users or by third parties but does not resemble a consultancy approach. This means that, even if funded by users, the evaluation process will not require direct or significant user involvement. For these reasons, we suggest that this model belongs on the producers/intermediaries side of the cost-sharing axis.

With respect to polarization, the realistic approach is rather agnostic. It is as likely to be used in low as in highly polarized contexts (especially if a third party funds the evaluation). For any evaluation to take place, it means minimally that gaining a better understanding of the program is a priority for someone. This does not mean, however, that there is consensus among stakeholders on this assessment. This tallies with Rogers' (1999) analysis, when she points out that realistic evaluation may encounter resistance from particular stakeholders if conclusions have implications that run contrary to their preferences or interests. For these reasons, we believe the realistic approach does not take any clear stand on polarization and thus occupies the lower half of the conceptual model (see Figure 2).

Empowerment evaluation (Fetterman and Wandersman)—Empowerment evaluation is a model that aims at improving the program evaluated, mostly through the development of the skills and autonomy of selected participants (Fetterman, 2005; Fetterman and Wandersman, 2005). It is generally conducted with a pool of program staff members and stakeholders who will be actively involved in identifying evaluation questions, collecting and analyzing data, and formulating next steps. Alkin and Christie (2004) place empowerment evaluation on the use branch of their tree, while Stufflebeam and Shinkfield (2007) consider it can easily turn into a form of pseudo-evaluation:

The empowerment activities move into the pseudoevaluation range when an external evaluator credits an internal evaluation as his own, credits a flawed internal evaluation as sound, stands silent when the client attributes the evaluation findings to the external evaluator, or fails to ensure that the evaluation will be subjected to an independent metaevaluation. (Stufflebeam and Shinkfield, 2007: 154)

The skill development component at the core of this model rests on intense and active stakeholder participation in, or even ownership of, the evaluation (Miller and Campbell, 2006). Empowerment evaluation thus clearly stands at the user pole of the cost-sharing axis (see Figure 3).

Regarding the polarization axis, Miller and Campbell (2006) surveyed 46 studies that identified themselves as empowerment evaluations. One of their core conclusions is the importance of working in a democratic/deliberative mode to help people analyze and improve their programs. In other words, to succeed, empowerment evaluations depend in large part on consensus-building. This requirement makes it hard to imagine an empowerment evaluation being conducted in a highly polarized context without it adding oil to the fire. We have therefore positioned this model in the upper left quadrant of our conceptual framework.

Democratic evaluation (House and Howe)—According to House and Howe, democratic evaluators are advocates of democracy (1999). Their aim is to provide society with reliable and relevant information to support public debates (MacDonald and Kushner, 2005). They should not be advocates for particular stakeholder groups, nor should they ‘play the role of neutral facilitators among advocates of competing “value summaries” or stakeholder “constructions”’ (House and Howe, 1999: 96). All perspectives should be represented in their evaluations, balancing out values and interests. This approach requires an independent evaluator position (MacDonald and Kushner, 2005). Clearly, in democratic evaluation, costs are not shared but are borne mainly by the evaluator, placing this model on the producers/intermediaries pole on the cost-sharing axis (Figure 4).

Likewise, the strong accent in this model on the need for a neutral evaluator rests on the assumption that stakeholders’ interests and points of view are divergent. However, democratic evaluation goes beyond the recognition that polarization matters and takes a clear stand on the fact that, in our societies, some are less equal than others. The evaluator’s social role is thus to encourage democratic debates and to contribute to the rebalancing of power relations in program-related decisions. As stated by MacDonald and Kushner, democratic evaluation is most appropriate in contexts where there are ‘widespread concerns’ (MacDonald and Kushner, 2005: 113). Democratic evaluation therefore belongs in the lower right quadrant of our conceptual framework.

Part III: Putting it all together

A first observation is that the two-dimensional space we developed to conceptualize use processes allows for effective contrasting of evaluation models. Each of the four models occupies a specific position in this space, and each, explicitly or not, is based on a particular conception of the relation between users and evaluators as well as on assumptions about the nature of the context. These models obviously rest on quite different assumptions, yet we believe the conceptual dimensions we developed are useful to analyze their similitude and divergences. Nevertheless, the actual potential of these conceptual dimensions goes further. Results from our systematic review of the literature suggest that both the level and the nature

of results use will vary according to specific positions on the cost-sharing and polarization axes.

First, as soon as users are willing to bear most of the costs, the potential for use increases dramatically. This observation is supported by converging evidence from all disciplinary traditions surveyed. Our data suggest there is a utilization ‘paradise’ in our two-dimensional space (see Figure 5). Second, there is also much robust evidence, mostly from political science (Bryant, 2003; Carpenter et al., 2004; Hall and Deardorff, 2006; Lowery et al., 2005; Phillips and Phillips, 1984; Smith, 1995), showing that high levels of polarization are associated with situations in which stakeholders are willing to invest significant amounts of resources and energy to get heard (Bourgeois and Nizet, 1993; Bryant, 2003; Heaney, 2006; Smith, 1995). This investment takes the form of lobby-like activities where ‘amenable’ evaluation results are likely to be used as ammunition in political battles. The literature offers quite convincing theoretical demonstrations as well as some empirical evidence that this translates into potentially significant levels of use. The lobbying zone will also intersect to some extent with the utilization paradise when users are ready to invest significant resources in a given evaluation with the purely tactical aim of supporting their own view against divergent information. Finally, there is a third zone we describe as the ‘knowledge-driven swamp’. Here we find low-polarized contexts in which the evaluator bears most of the costs and which are associated with very little potential for use. Results produced in this quadrant are likely to join the ever-growing pile of ignored advice and shelved reports.

Parenthetically, readers may have noted from the adjectives describing the quadrants in our framework that we consider that even strategic (Whiteman, 1985) or symbolic (Knorr, 1977) use (using evaluation results to support pre-existing positions) is probably preferable to the absence of any use. This self-confessed bias rests on the idea that even the most partisan use increases information circulation and that, as long as the information is not totally erroneous, this is better than nothing. Practically, notwithstanding how objective evaluation results may be, their use will take place in political systems (Chelimsky, 1987; Palumbo, 1987; Weiss, 1987b). Greenberg and Mandell (1991) provide a good discussion of empirical data supporting the view that strategic use of given selective information contributes to its diffusion within the system and increases the likeliness of its uptake by other, more neutral users. However, some may also offer coherent arguments supporting the opposite view – that no use is preferable to partisan use. For those taking this second perspective, it is important to note that the results of our systematic review suggest that most use occurring in polarized contexts (the two right quadrants of the graph) is likely to be selective and partisan, and even more so as polarization grows.

Figure 5 illustrates the three zones of use and reproduces our interpretation of our four illustrative models. This figure shows that the model we developed includes hypotheses on each evaluation model’s probability of fostering use.

Debates on results use have traditionally focused on the means an evaluator can or should employ to influence and encourage use. The evaluation process is generally targeted as an avenue for the evaluator to increase evaluation results use. Even if evaluation models are based on contrasted epistemological and methodological foundations, fostering use is

generally thought to depend on the evaluator's abilities to strategically harness the evaluation process and exploit his personal communication qualities. Yet, our analysis shows it would be a mistake to think results use depends primarily on the model used or the evaluator's qualities. Certainly these two factors play a role, but focusing on them to explain results use can obscure an essential determinant: the evaluation context. The results discussed in the first section of this article suggest that contextual characteristics explain in large part the level and nature of results use. In Part II, we positioned selected evaluation models in the two-dimensional framework, characterizing context according to those models' core components. In this section, we discuss the meaning of this alignment between context characteristics and models and the implications this can have for evaluation practice.

It is important to recognize that implicit hypotheses regarding the level of freedom evaluators actually have in selecting the model used in each context have different implications for practice. At the macro level, institutional rules and requirements, such as governmental frameworks, might strongly restrict evaluators' capacity to choose what evaluation model they will use. At the micro level, the choice of evaluation model is likely to be structured by evaluators' preferences, skills and training, as well as by their status (consultant, researcher, bureaucrat, etc.), which determines their degree of independence. Depending on where we position the evaluator's actual capacity to choose the evaluation model implemented, three different conclusions can be drawn about the fit between model and context. These are summarized in Table 2.

The first case is what we call a contingency theory driven fit. If evaluators are free to rely on the evaluation model they want and are skilful in diagnosing the context's characteristics, then good evaluators will be able to choose models that fit with the context. The last case is what we call a natural selection fit. If evaluators have strong and stable preferences regarding model choice and are sensitive to the contextual appropriateness of their favoured model, then they will tend to specialize in specific niche contexts where the model they work with is a good fit. Between these two outcomes where models and contexts are in fit, there are possible outcomes where the fit between model and context will occur randomly and sometimes result in misfit. It is important to stress that the result of such a misfit between the evaluation model used and the context characteristics is not only that little use of the evaluation result can be expected, but more centrally, that it will prove difficult or even impossible to implement the evaluation model itself in a coherent way.

In the case of a contingency theory driven fit, it should be stressed that evaluators will have to conclude from time to time that they are working in a context where a high level of instrumental use is unlikely. If we are right in our premise that characteristics of the evaluand's context explain the nature and the level of use, over and above the choice of evaluation model, then, short of influencing the characteristics of the context itself, evaluators' capacity to increase use is effectively limited.

In cases of natural selection fit, some evaluators will specialize in contexts propitious to high levels of instrumental use, while others will have a harder time getting their results used. This is likely to produce rather different views about the practice of evaluation. It is easy to see parallels between this idea and the deep-seated divergent views of Patton and Weiss over

the evaluator's capacity to increase use, drawn from each of their own considerable experiences of evaluation practice (Patton, 1988; Weiss, 1987a, 1988).

Conclusion

This article's contribution lies in providing a framework for understanding utilization and predicting the impact of evaluation results. A systematic review of the literature by Contandriopoulos et al., (2010) demonstrated the determining influence of contextual characteristics on the use of research results. Our model goes a step further by establishing the links between the primary components of models and their positioning in relation to the dimensions influencing results use: the degree of polarization and the distribution of costs between producers and users. Thus, evaluation models advance certain constituent principles and hypotheses that have the effect of placing them in different contexts. The correspondence between what the literature tells us about the influence of contextual characteristics and the models' positions as self-determined by their core components leads to the notion of fit, as discussed above. In conclusion, we would like to highlight two implications of this discussion for evaluation practice.

First, the desire to maximize results use should not lead systematically to selecting evaluation models that are more oriented toward use, nor to eliminating certain evaluation models that might be found in contexts where instrumental use is less probable. Doing so would involve deliberately contributing to a natural selection fit by deeming unsuitable for evaluation purposes contexts where easy instrumental use is unlikely. Our own view of the evaluative endeavor is broader, being based on the principle that appropriate models exist for any and all contexts and that while a context-driven low potential for use should be taken into account, it does not constitute a reason not to evaluate a program.

Second – and this point would deserve a paper of its own – it could also be argued that context characteristics are not immutable and that astute evaluators might indeed influence them. One avenue would be to influence the context and eventually make it more favorable for use by acting on one of the axes, either by trying to reduce polarization or by trying to change the representations of potential users to promote their participation. Going down this road means abandoning the simple role of evaluator and adopting a politically involved or lobbying role. As well, following that path, if not a utopian proposition, will certainly be a long and demanding one.

Acknowledgments

Funding

A Canadian Institutes for Health Research (CIHR) open grant (MOP 84259) made this research possible.

References

- Ainsworth S. Regulating lobbyists and interest group influence. *Journal of Politics*. 1993; 55(1):41–56.
- Albaek E. Between knowledge and power: utilization of social science in public policy making. *Policy Sciences*. 1995; 28(1):79–100.

- Alkin, MC., Christie, CA. An evaluation theory tree. In: Alkin, MC., editor. *Evaluation Roots. Tracing Theorists' Views and Influences*. Thousand Oaks, CA: SAGE; 2004. p. 12-65.
- Amara N, Ouimet M, Landry R. New evidence on instrumental, conceptual, and symbolic utilization of university research in government agencies. *Science Communication*. 2004; 26(1):75–106.
- Arendt, H. *La crise de la culture*. Paris: Folio; 1972. Chapitre VII: Vérité et politique; p. 289-336.
- Austen-Smith D, Wright JR. Competitive lobbying for a legislator's vote. *Social Choice and Welfare*. 1992; 9(3):229–57.
- Bardach E. The dissemination of policy research to policymakers. *Knowledge-Creation Diffusion Utilization*. 1984; 6(2):125–44.
- Baumgartner F, Leech BL. The multiple ambiguities of 'counteractive lobbying'. *American Journal of Political Science*. 1996; 40(2):521–42.
- Beyer JM, Trice HM. The utilization process: a conceptual framework and synthesis of empirical findings. *Administrative Science Quarterly*. 1982; 27(4):591–622.
- Black N. Evidence based policy: proceed with care. *British Medical Journal*. 2001; 323:275. [PubMed: 11485961]
- Booth T. Researching policy research: issues of utilization in decision making. *Knowledge-Creation Diffusion Utilization*. 1990; 12(1):80–100.
- Bourgeois E, Nizet J. Influence in academic decision-making: towards a typology of strategies. *Higher Education*. 1993; 26(4):387–409.
- Bowen S, Zwi AB. Pathways to 'evidence-informed' policy and practice: a framework for action. *PLoS Medicine*. 2005; 2(7):600–5.
- Brunsson N. The irrationality of action and action rationality: decisions, ideologies and organizational actions. *Journal of Management Studies*. 1982; 19(1):29–44.
- Bryant T. A critical examination of the hospital restructuring process in Ontario, Canada. *Health Policy*. 2003; 64(2):193–205. [PubMed: 12694955]
- Burstein P, Hirsh CE. Interest organizations, information, and policy innovation in the US Congress. *Sociological Forum*. 2007; 22(2):174–99.
- Campbell D, Redman S, Jorm L, Cooke M, Zwi AB, Rychetnik L. Increasing the use of evidence in health policy: practice and views of policy makers and researchers. *Australia and New Zealand Health Policy*. 2009; 6(1):21. [PubMed: 19698186]
- Caplan N. The two-communities theory and knowledge utilization. *The American Behavioral Scientist*. 1979; 22(3):459–70.
- Caplan, N., Morisson, A., Stambaugh, RJ. *The Use of Social Science Knowledge in Policy Decisions at the National Level: A Report to Respondents*. Ann Arbor, MI: Center for Research on Utilization of Scientific Knowledge (The University of Michigan); 1975.
- Carpenter DP, Esterling KM, Lazer DMJ. The strength of strong ties: a model of contact-making in policy networks with evidence from U.S. health politics. *Rationality and Society*. 2003; 15(4):411–40.
- Carpenter DP, Esterling KM, Lazer DMJ. Friends, brokers, and transitivity: who informs whom in Washington politics? *Journal of Politics*. 2004; 66(1):224–46.
- Chelimsky, E. The politics of program evaluation. In: Cordray, DS, Bloom, HS., Light, JR., editors. *Evaluation Practice in Review*. San Francisco, CA: Josey-Bass; 1987. p. 5-21.
- Coglianesi C, Zeckhauser R, Parson E. Seeking truth for power: informational strategy and regulatory policymaking. *Minnesota Law Review*. 2004; 89(2):277–341.
- Contandriopoulos D, Lemire M, Denis J-L, Tremblay É. Knowledge exchange processes in organizations and policy arenas: a narrative systematic review of the literature. *Milbank Quarterly*. 2010; 88(4):444–83. [PubMed: 21166865]
- Cousins JB, Leithwood KA. Current empirical research on evaluation utilization. *Review of Educational Research*. 1986; 56(3):331–64.
- Dobrow MJ, Goel V, Upshur REG. Evidence-based health policy: context and utilisation. *Social Science & Medicine*. 2004; 58(1):207–17. [PubMed: 14572932]

- Elliott H, Popay J. How are policy makers using evidence? Models of research utilisation and local NHS policy making. *Journal of Epidemiology and Community Health*. 2000; 54(6):461–8. [PubMed: 10818123]
- Epstein D, Ohalloran S. A theory of strategic oversight: Congress, lobbyists, and the bureaucracy. *Journal of Law Economics & Organization*. 1995; 11(2):227–55.
- Fetterman, DM. Empowerment evaluation. In: Mathison, S., editor. *Encyclopedia of Evaluation*. Thousand Oaks, CA: SAGE; 2005. p. 125–9.
- Fetterman, DM., Wandersman, A. *Empowerment Evaluation. Principles in Practice*. New York: The Guilford Press; 2005.
- Florio E, Demartini JR. The use of information by policy-makers at the local community level. *Knowledge: Creation Diffusion Utilization*. 1993; 15(1):106–23.
- Forbes A, Griffiths P. Methodological strategies for the identification and synthesis of ‘evidence’ to support decision-making in relation to complex healthcare systems and practices. *Nursing Inquiry*. 2002; 9(3):141–55. [PubMed: 12199878]
- Freeman R. Epistemological bricolage: how practitioners make sense of learning. *Administration & Society*. 2007; 39(4):476–96.
- Gabbay J, Le May A, Jefferson H, Webb D, Lovelock R, Powell J, et al. A case study of knowledge management in multiagency consumer-informed ‘communities of practice’: implications for evidence-based policy development in health and social services. *Health: An Interdisciplinary Journal for the Social Study of Health, Illness and Medicine*. 2003; 7(3):283–310.
- Greenberg DH, Mandell MB. Research utilization in policymaking: a tale of two series (of social experiments). *Journal of Policy Analysis and Management*. 1991; 10(4):633–56.
- Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of Innovations in Service Organizations: Systematic Review and Recommendations. *The Milbank Quarterly*. 2004; 82(4): 581–629. [PubMed: 15595944]
- Haas P. Introduction: epistemic communities and international policy coordination. *International Organization*. 1992; 46(1):1–35.
- Hall RL, Deardorff AV. Lobbying as legislative subsidy. *American Political Science Review*. 2006; 100(1):69–84.
- Harries U, Elliott H, Higgins A. Evidence-based policy-making in the NHS: exploring the interface between research and the commissioning process. *Journal of Public Health Medicine*. 1999; 21(1): 29–36. [PubMed: 10321856]
- Havelock, RG. *Planning for Innovation through Dissemination and Utilization of Knowledge*. Ann Arbor, MI: Institute for Social Research (University of Michigan); 1969.
- Heaney MT. Brokering health policy: coalitions, parties, and interest group influence. *Journal of Health Politics Policy and Law*. 2006; 31(5):887–944.
- Heclo, H. Issue networks and the executive establishment: government growth in an age of improvement. In: King, A., editor. *The New American Political System*. Washington, DC: American Enterprise Institute; 1978. p. 87–124.
- Heinz, JP., Laumann, EO., Nelson, RL., Salisbury, RH. *The Hollow Core: Private Interests in National Policy Making*. Cambridge, MA: Harvard University Press; 1993.
- Henry GT, Mark MM. Beyond use: understanding evaluation’s influence on attitudes and actions. *American Journal of Evaluation*. 2003; 24(3):293–314.
- House, ER., Howe, KR. *Values in Evaluation and Social Research*. Thousand Oaks, CA: SAGE; 1999.
- Huberman M. Steps toward an integrated model of research utilization. *Knowledge: Creation Diffusion Utilization*. 1987; 8(4):586–611.
- Huberman M. Research utilization: the state of the art. *Knowledge, Technology & Policy*. 1994; 7(4): 13–33.
- Jacobson N, Butterill D, Goering P. Consulting as a strategy for knowledge transfer. *Milbank Quarterly*. 2005; 83(2):299–321. [PubMed: 15960773]
- Jordan G, Maloney WA. Accounting for subgovernments – Explaining the persistence of policy communities. *Administration & Society*. 1997; 29(5):557–83.

- Kickert, WJM, Klijn, E-H., Koppenjan, JFM., editors. *Managing Complex Networks*. London: SAGE; 1999.
- Klijn EH. Analyzing and managing policy processes in complex networks: a theoretical examination of the concept policy network and its problems. *Administration & Society*. 1996; 28(1):90–119.
- Knorr, KD. Policy makers' use of social science knowledge: symbolic or instrumental?. In: Weiss, CH., editor. *Using Social Research in Public Policy Making*. Lexington, MA: Lexington Books; 1977. p. 165–82.
- Knott J, Wildavsky A. If dissemination is the solution, what is the problem? *Knowledge-Creation Diffusion Utilization*. 1980; 1(4):537–78.
- Landry R, Amara N, Lamari M. Climbing the ladder of research utilization – Evidence from social science research. *Science Communication*. 2001; 22(4):396–422.
- Landry R, Lamari M, Amara N. The extent and determinants of the utilization of university research in government agencies. *Public Administration Review*. 2003; 63(2):192–205.
- Larocca R. Strategic diversion in political communication. *Journal of Politics*. 2004; 66(2):469–91.
- Lavis JN, Robertson D, Woodside JM, McLeod CB, Abelson J. How can research organizations more effectively transfer research knowledge to decision makers? *Milbank Quarterly*. 2003; 81(2):221–48. [PubMed: 12841049]
- Lemieux, V. *Les coalitions: Liens, transactions et contrôles*. Paris: P.U.F; 1998.
- Lindblom CE. The science of 'muddling through'. *Public Administration Review*. 1959; 19(2):79–88.
- Lindblom, CE., Cohen, DK. *Usable Knowledge: Social Science and Social Problem Solving*. New Haven, CT: Yale University Press; 1979.
- Lomas J. Finding audiences, changing beliefs: the structure of research use in Canadian health policy. *Journal of Health Politics, Policy and Law*. 1990; 15(3):525–42.
- Lowery D, Gray V, Wolak J, Godwin E, Kilburn W. Reconsidering the counter-mobilization hypothesis: health policy lobbying in the American states. *Political Behavior*. 2005; 27(2):99–132.
- Lynn, LE., editor. *Knowledge and Policy: The Uncertain Connection*. Washington, DC: National Academy of Sciences; 1978.
- MacDonald, B., Kushner, S. Democratic evaluation. In: Mathison, S., editor. *Encyclopedia of Evaluation*. Thousand Oaks, CA: SAGE; 2005. p. 109–13.
- Majone, G. *Evidence, Argument and Persuasion in the Policy Process*. New Haven, CT: Yale University Press; 1989.
- March, JG. *Decisions and Organizations*. New York: Blackwell; 1988.
- March, JG., Olsen, JP. *Ambiguity and Choice in Organizations*. Bergen: Universitetsforlaget; 1976.
- McNie EC. Reconciling the supply of scientific information with user demands: an analysis of the problem and review of the literature. *Environmental Science & Policy*. 2007; 10(1):17–38.
- Milbrath LW. Lobbying as a communication process. *Public Opinion Quarterly*. 1960; 24(1):32–53.
- Milbrath, LW. *The Washington Lobbyists*. Chicago, IL: Rand McNally and Company; 1963.
- Miller RL, Campbell R. Taking stock of empowerment evaluation: an empirical review. *The American Journal of Evaluation*. 2006; 27(3):296–319.
- Mitton C, Adair CE, McKenzie E, Patten SB, Perry BW. Knowledge transfer and exchange: review and synthesis of the literature. *The Milbank Quarterly*. 2007; 85(4):729–68. [PubMed: 18070335]
- Nonaka I. A dynamic theory of organizational knowledge creation. *Organization Science*. 1994; 5(1):14–37.
- Olson, M. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press; 1965.
- Palumbo, DJ. Politics and evaluation. In: Palumbo, DJ., editor. *The Politics of Program Evaluation*. Newbury Park, CA: SAGE; 1987. p. 12–46.
- Patton MQ. The evaluator's responsibility for utilization. *American Journal of Evaluation*. 1988; 9(5):5–24.
- Patton, MQ. *Utilization-Focused Evaluation*. Thousand Oaks, CA: SAGE; 1997.
- Patton, MQ. Utilization-focused evaluation. In: Mathison, S., editor. *Encyclopedia of Evaluation*. Thousand Oaks, CA: SAGE; 2005. p. 429–32.

- Pawson, R., Tilley, N. *Realistic Evaluation*. London: SAGE; 1997.
- Pawson, R., Tilley, N. *Realistic evaluation*. In: Mathison, S., editor. *Encyclopedia of Evaluation*. Thousand Oaks, CA: SAGE; 2005. p. 362-7.
- Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review – a new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy*. 2005; 10(suppl 1):21–34. [PubMed: 16053581]
- Perelman, C., Olbrechts-Tyteca, L. *The New Rhetoric: A Treatise on Argumentation*. Notre Dame: University of Notre Dame Press; 1969.
- Peterson, MA. Intensive care: how Congress shapes health policy. In: Mann, TE., Ornstein, NJ., editors. *Intensive Care: How Congress Shapes Health Policy*. Washington, DC: American Enterprise Institute; 1995. p. 79-125.
- Phillips KB, Phillips LT. Research, politics, and the dynamics of policy development: a case study of motor carrier regulatory reform. *Policy Sciences*. 1984; 17(4):367–84.
- Pressman, J., Wildavsky, A. *Implementation*. Berkeley: University of California Press; 1973.
- Rhodes R. Policy networks: a British perspective. *Journal of Theoretical Politics*. 1990; 2(3):293–317.
- Rich RF. The pursuit of knowledge. *Knowledge-Creation Diffusion Utilization*. 1979; 1(1):6–30.
- Rogers PJ. Book review: *Realistic Evaluation*. *The American Journal of Evaluation*. 1999; 20(2):381–3.
- Russell J, Greenhalgh T, Byrne E, McDonnell J. Recognizing rhetoric in health care policy analysis. *Journal of Health Services Research & Policy*. 2008; 13(1):40–6.
- Sabatier P. The acquisition and utilization of technical information by administrative agencies. *Administrative Science Quarterly*. 1978; 23(3):396–417. [PubMed: 10297211]
- Salisbury RH, Heinz JP, Laumann EO, Nelson RL. Who works with whom? Interest group alliances and opposition. *American Political Science Review*. 1987; 81(4):1217–34.
- Shadish, WR., Cook, TD., Leviton, LC. Good theory for social program evaluation. In: Shadish, WR., Cook, TD., Leviton, LC., editors. *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA: SAGE; 1991. p. 36-67.
- Shulha LM, Cousins JB. Evaluation use: theory, research, and practice since 1986. *Evaluation Practice*. 1997; 18(3):195–208.
- Simon, HA. Designing organizations for an information-rich world. In: Greenberger, M., editor. *Computers, Communication, and the Public Interest*. Baltimore, MD: The Johns Hopkins Press; 1971. p. 37-72.
- Sloof R, Van Winden F. Show them your teeth first! A game-theoretic analysis of lobbying and pressure. *Public Choice*. 2000; 104(1–2):81–120.
- Smith RA. Advocacy, interpretation, and influence in the U.S. Congress. *The American Political Science Review*. 1984; 78(1):44–63.
- Smith RA. Interest group Influence in the U.S. Congress. *Legislative Studies Quarterly*. 1995; 20(1): 89–139.
- Stufflebeam, DL., Shinkfield, AJ. The nature of program evaluation theory. In: Stufflebeam, DL., Shinkfield, AJ., editors. *Evaluation. Theory, Models and Applications*. San Francisco, CA: Jossey-Bass; 2007. p. 57-79.
- Van de Ven AH, Schomaker MS. Commentary: the rhetoric of evidence-based medicine. *Health Care Management Review*. 2002; 27(3):89–91. [PubMed: 12146787]
- Weick, KE. *Sensemaking in Organizations*. Thousand Oaks, CA: SAGE; 1995.
- Weiss, CH. Introduction. In: Weiss, CH., editor. *Using Social Research in Public Policy Making*. Lexington, MA: Lexington Books; 1977a. p. 1-22.
- Weiss CH. Research for policy's sake: the enlightenment function of social research. *Policy Analysis*. 1977b; 3(4):531–45.
- Weiss, CH., editor. *Using Social Research in Public Policy Making*. Lexington, MA: Lexington Books; 1977c.
- Weiss, CH. Improving the linkage between social research and public policy. In: Lynn, LE., editor. *Knowledge and Policy: The Uncertain Connection*. Washington, DC: National Academy of Sciences; 1978. p. 23-81.

- Weiss CH. The many meanings of research utilization. *Public Administration Review*. 1979; 39(5): 426–31.
- Weiss, CH. Ideology, interests and information: the basis of policy positions. In: Callahan, D., Jennings, B., editors. *Ethics, the Social Sciences and Policy Analysis*. New York: Plenum Press; 1983. p. 213–45.
- Weiss CH. Evaluation for decisions: is anybody there? Does anybody care? *American Journal of Evaluation*. 1987a; 9(5):5–19.
- Weiss, CH. Where politics and evaluation research meet. In: Palumbo, DJ., editor. *The Politics of Program Evaluation*. Newbury Park, CA: SAGE; 1987b. p. 47–70.
- Weiss CH. If program decisions hinged only on information: a response to Patton. *American Journal of Evaluation*. 1988; 9(15):15–28.
- Weiss, CH., Bucuvalas, MJ. *Social Science Research and Decision-Making*. Lexington, MA: Lexington Books; 1980a.
- Weiss CH, Bucuvalas MJ. Truth tests and utility tests: decision-makers' frames of reference for social science research. *American Sociological Review*. 1980b; 45(2):302–13.
- Whiteman D. Reaffirming the importance of strategic use: a two-dimensional perspective on policy analysis in Congress. *Knowledge: Creation, Diffusion, Utilization*. 1985; 6(3):203–24.
- Wright JR. Contributions, lobbying, and committee voting in the U.S. House of Representatives. *The American Political Science Review*. 1990; 84(2):417–38.

Biographies

Damien Contandriopoulos, PhD, is Associate Professor at the Faculty of Nursing and researcher at the Public Health Research Institute (IRSPUM) of the University of Montreal. Please address correspondence to: C.P. 6128 succursale Centre-ville, Montréal Québec H3C 3J7, Canada. [email: damien.contandriopoulos@umontreal.ca]

Astrid Brousselle, PhD, is Associate Professor at the Department of Community Health Sciences, University of Sherbrooke, and researcher at the Charles-LeMoyne Hospital Research Center. Please address correspondence to: CR-HCLM- Université de Sherbrooke, Campus de Longueuil, 150 place Charles LeMoyne, bureau 200, CP 11, Longueuil (Québec) J4K-0A8, Canada. [email: astrid.brousselle@usherbrooke.ca]

Figure 1.
Positioning of Patton's utilization-focused evaluation model

Figure 2.
Positioning of Pawson's and Tilley's realistic evaluation model

Figure 3.
Positioning of Fetterman et al.'s empowerment evaluation model

Figure 4.
Positioning of House's and Howe's democratic evaluation model

Figure 5.
Synthesis of models and hypotheses on the nature of use

Table 1

Selected contrasted evaluation models

Models	Classification according to...	
	Alkin & Christie (2004)	Stufflebeam and Shinkfield (2007)
Utilization-focused evaluation/Patton	Use	Eclectic approaches
Realistic evaluation/Pawson & Tilley	Methods	Improvement/accountability approaches
Empowerment evaluation/Fetterman & Wandersman	Use	Pseudoevaluations
Democratic evaluation/House & Howe	Valuing	Social agenda/advocacy approaches

Table 2

Summary of the relation between freedom in model choice and fit between model and context

Hypotheses about the determinants explaining the evaluation model used		Level of expected fit between the characteristics of the evaluation model used and the context
First level hypothesis	Second level hypothesis	
Evaluators can choose the evaluation model they will use	(Good) Evaluators will be able to choose the proper models for each context	Contingency theory driven fit
	(Bad) Evaluators will fail at properly choosing their model	Random fit
Evaluators can't/won't choose the evaluation model they will use	Institutional characteristics will structure model choice	
	<div>Evaluators have strong and stable preferences regarding models</div> <div>(Bad) Evaluators will try to use their favourite model whatever the context</div> <div>Model preferences will push evaluators to work in contexts appropriate to their favorite models</div>	Natural selection fit