



Universiteit
Leiden
The Netherlands

The politics of evaluation

Mavrot, C.; Pattyn, V.E.; Ladner, A.; Sager, F.

Citation

Mavrot, C., & Pattyn, V. E. (2022). The politics of evaluation. In A. Ladner & F. Sager (Eds.), *Handbook on the politics of public administration* (pp. 244-255). Edward Elgar.
doi:10.4337/9781839109447.00029

Version: Accepted Manuscript

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/3620642>

Note: To cite this publication please use the final published version (if applicable).

This is the pre-print version of the article. Please refer to:

Mavrot Céline, Pattyn Valérie (2022). “The Politics of Evaluation”, in Ladner Andreas, Sager Fritz, *Handbook on the Politics of Public Administration* (pp. 243-254), Cheltenham: Edward Elgar. ISBN: 978 1 83910 943

<https://www.e-elgar.com/shop/gbp/handbook-on-the-politics-of-public-administration-9781839109430.html>

Chapter 22 - The Politics of Evaluation

Céline Mavrot¹, Valérie Pattyn²

¹Yale University School of Public Health, USA, and University of Bern, Switzerland

²Leiden University, Institute of Public Administration, The Netherlands, and KU Leuven Public Governance Institute, Belgium

Abstract

This chapter focuses on the evaluation of public policy and administration. Institutionalized for decades in many places worldwide, policy evaluation has become a routinized and professionalized activity. However, assessing the performance of public agencies and the impact of the policies they implement is a highly political endeavor. As part of the executive branch and directly subordinate to governments, public agencies can be scrutinized and evaluated as an extension of political struggles. On the other hand, public agencies also use evaluation to advance their own objectives. This chapter examines the issues attached to the evaluation of public policies and administrative activity in relation to power games within and across the branches of government. It reveals how politicians and public servants can make strategic use of policy evaluations, as well as how this instrument serves not only reflexive and oversight purposes, but also agenda-setting ambitions. The chapter then presents an overview of the controlling, defensive and proactive functions of evaluation in policy struggles. Drawing upon these developments, the chapter underlines just how far from neutral evaluations can be in the politico-administrative game.

Introduction

As an activity aimed at assessing the quality and achievements of public action, policy evaluation has become an integral part of the political game in many systems. The formal features of evaluation make it an institutionalized element of the policy cycle, intended to provide an independent analysis of public programmes through established methods and procedures. However, there is much more than that to evaluation. Policy evaluations bring together a whole range of vested interests in given policy subsectors, and therefore hold a consequent arbitration power between concurrent policy narratives. Indeed, if policy evaluations rely on well-established scientific procedures, they are also based on “disparate bodies of knowledge” which serve as “multiple sets of evidence that inform and influence policy” (Head, 2008: 4). In other words, “there is not one evidence-base but several bases” (Head, 2008: 7).

Evaluations make a synthesis of these different sets of evidence through a series of categorizations that culminate in the public assessment of political action, which gives them a high salience for office holders. Moreover, while evaluations might be activated by various kind of institutional factors such as legal obligations, financial requirements or planning and control cycles, it is also not unusual that they are commissioned in the wake of crises, political incidents or disruptive events (Bovens et al., 2008). Evaluations are thus often found at the heart of politics.

Not only is evaluation located within the political game, but every step of the evaluation process is open to politicization, starting with the very decision to put a public policy under scrutiny. Moreover, since public action is multidimensional in nature, determining the criteria for assessing a policy is highly prone to interpretative dispute. Calls have been made to adopt a wider perspective on policy performance and to go beyond the classical rationalistic criteria of effectiveness and efficiency in favor of novel governance principles able to meet contemporary challenges such as environmental problems or transition issues (Kunseler & Vasileiadou, 2016). Not only the substance of evaluation criteria but also the procedure through which they are defined can further be open to confrontation. Determining which stakeholder deserves a seat at the table at each stage of the process is an intrinsically political question (Weaver & Cousin, 2004). There is a continuum between expert evaluations relying on top-down processes and result-based legitimacy, and participative evaluations emphasizing a bottom-up approach for a stronger process-based legitimacy (Sager & Mavrot, 2021). Finally, the use of evaluation results later in the policy cycle also opens the door to vivid debates between stakeholders including the officials in charge, the opposition or the public at large.

After briefly introducing policy evaluation and its role within the policy process, this chapter examines the many ways that evaluation is a part of policy disputes, with a specific focus on the use of evaluations by politicians and civil servants. It provides an overview of the variety of evaluation uses within the policy game, highlighting its controlling, defensive and proactive functions. It finally reflects on the role of evaluation as a democratic instrument, inevitably caught up in politicization processes but also able to provide a distinctive contribution to the public debate.

Public Policy and Evaluation

This section provides a short overview of the historical developments of policy evaluations and of their roles in the formulation and the reform of public policies. Policy evaluation comes in many guises, and different evaluators often have their own understanding of the essence of evaluation. Inspired by Scriven (Scriven, 1991: 139), in this chapter, we conceive of policy evaluation as a systematic assessment of a certain policy (or part of a policy), aimed at determining its merit or worth or value on the basis of certain criteria. We apply a comprehensive perspective to policy evaluation, and do not limit it to the last stage of the policy cycle (typically referred to as *ex post* evaluation). Policy evaluation

also has the potential to affect the design of policy and legislation (ex ante evaluation) or to inform decisions when implementing policies or legislation (coined as ex durante/ad interim evaluations).

In recent decades, policy evaluation has managed to secure a virtual sacred status (Dahler-Larsen, 2011: 3) as a key ingredient of good governance. A concise review of the chronology of evaluation dissemination helps in understanding the value attached to evaluation, and how different emphases have been placed across time. Different waves of evaluation diffusion can be distinguished, each posing specific questions for evaluation (Furubo & Sandahl, 2002 ; Stame, 2003; Vedung, 2010), and bringing different methods to the evaluation toolbox (Francesco & Pattyn, 2021).

The first wave emerged in the 1960s-1970s, with the first programs of the Great Society launched in the US. Evaluation requirements were systematically incorporated into the introduction of new social programmes, and reflected a political culture favoring a strong rationalist and engineering approach to policy making. Evert Vedung (2010) labelled the 1960s as the '*scientific wave*', with evaluations (preferably randomized controlled trials) being deployed to determine whether social programmes were effective in achieving the intended results.

The importance attached to scientific assessment was heavily criticized in the early 1970s, however, with scholars calling attention to the perspective of the users and practitioners on public service and policy quality when evaluating policy. Accordingly, this gave impetus to the emergence of a second wave of more democratic, responsive and participatory types of evaluation (Vedung, 2010), under the flag of '*dialogic evaluation*'.

A subsequent, third, wave was launched with the introduction of the New Public Management (NPM) doctrine in the late 1970s by countries of the Anglo-Saxon tradition, and some Northern European countries. With NPM, and the associated incorporation of private sector principles in the public sector (such as deregulation, privatization, agentification), the concept of '*value for money*' conquered a central position. Accordingly, this '*neo-liberal wave*' put an emphasis on efficiency-oriented (such as cost-benefit analysis; cost-effectiveness analysis) and client-oriented approaches to evaluation (Vedung, 2010). With the adoption of NPM, a key distinction was also introduced between accountability and learning as guiding rationales for policy evaluation. As Nicoletta Stame put it (2003: 38) the learning function draws attention to the ability of actors and stakeholders to benefit from the evaluation. Besides, with the clear identification of executive agents and leading principals, the former can be held more accountable for his/her actions toward the latter.

The most recent fourth wave came with the expansion of evaluation activity and capacity across Europe. The European Structural Funds have played a major role in this regard, with monitoring and evaluation requirements linked to the major investments in social funds (Stame, 2003). Besides, the 2008 financial crisis pushed evaluation high on the political agenda. Evaluation had the potential to support the austerity shift, to help governments making decisions about what to cut and to find new ways of addressing public needs and delivering services (Stame, 2012). The renaissance of the scientific '*evidence-based policy*' wave (Vedung, 2010), and the associated '*what works*' discourse is equally situated in this fourth wave. Importantly, each of these waves has left its sediments until today. The diversity and richness of present-day evaluation practice and the evaluators' toolbox of methods are precisely the result of the traces that these different waves have left.

While evaluations are research-based by nature and rely on scientific procedures, it should be clear that such assessment does not need to be conducted by scientists per se. Evaluation is an activity par excellence in which a wide (and ever growing) range of actors is active, including civil society organizations, courts of audit, think tanks, consultants, academics, Members of Parliament (MPs), and civil servants in government departments, to name just a few. Yet, depending on one's approach to evaluation, and a country's knowledge regime, a specific configuration of actors can come to the fore (see also Francesco & Pattyn, 2021). For instance, when adopting an evidence-based stance toward evaluation, academics can have a key role, just as independent government research institutes can.

Consultants are often mobilized when a more neoliberal approach to evaluation is applied. In a more responsive evaluation tradition in turn, it is common practice to involve citizens affected by the policy being evaluated, and ask their input in designing and executing the evaluation.

The Politics of Evaluation

The following section takes a closer look at the role of evaluations within the policy game. It first provides some reflections on the fundamentally political nature of policy evaluations, before turning to the potential strategic use of policy evaluations by political and administrative actors during their policy formulation and policy implementation activities.

Evaluation at the Heart of Policy Struggles

As should be clear from the concise historic overview above, policy evaluations are important, albeit often overlooked, policy-making instruments (Stephenson et al., 2019). As an ensemble of specialized procedures aimed at performing diagnostics on public actions, they are part of the policy narrative, which raises important stakes around them. Two general conceptions have emerged in the literature along with the historical development of the policy evaluation field. From the first perspective, evaluations are seen as merely neutral instruments that measure the achievement of objectives as previously defined in policy programmes; the second perspective, however, brings politics back into evaluations and acknowledges that they are intrinsically linked to power frameworks (McConnell, 2010: 347). These two distinct conceptions can be considered under the labels of a rationalistic tradition emphasizing value neutrality and an argumentative tradition, apprehending evaluation as inherently marked by confrontation (Bovens et al., 2008). Following the second perspective, we can conceptualize evaluation as “both a normative exercise, in that it presumes standards against which performance will be assessed, and a political exercise, in that attaching certain labels to a programme or project can have significant consequences for those involved in and affected by it” (Bovens & ‘t Hart, 2016: 655). Hence, policy evaluations implicate power relationships at various steps of the process, notably in defining evaluation standards against which to operationalize an assessment, framing a policy as a failure or a success and proposing corrective actions that could affect the future of the policy field and its stakeholders.

After all, depending on the values and the related evaluation criteria placed centrally in a policy evaluation, the assessment findings can differ greatly (Stephenson et al., 2019; see also Peters et al., 2018: 3). Those stakeholders managing to have their assessment criteria considered can, as such, directly or indirectly influence the eventual policy decision. These power plays inherent to an evaluation are often not explicitly examined. Moreover, with evaluations often involving a wide range of stakeholders, and ‘serving many masters’, interests do not always coincide (Palumbo & Hallett, 1993). It is an important feature of evaluations that they take a stance in matters that concern a wide range of actors with different types of involvement in the policy. Implementing actors might pursue various or contradictory objectives; implementation partners from different backgrounds might be attached to specific professional norms; various policy stakeholder might be differently affected by the policy; and the final target groups might approve, disregard or oppose the policy.

As such, policy evaluation is a social activity that cannot be reduced to a series of value-free measuring operations. It also engages competing reputations, narratives and interests (Bovens & ‘t Hart, 2016). Hence, policy evaluation is firmly embedded in politics, to say the least, which explains the sometimes high level of conflict that surrounds it. In most liberal democracies, evaluation enjoys a certain degree of institutionalization (Jacob et al., 2015; Stockmann et al., 2020a), whether as an instrument of

political control over the administration, a self-assessment tool used by politico-administrative actors to guide their actions or a means of democratic accountability toward the public. This makes evaluation an important source of control, along with other traditional democratic surveillance bodies such as the media or justice courts. On paper, and as outlined above, policy evaluation has some overarching purposes that make it a useful instrument for governance procedures: fostering public accountability, promoting policy learning for the betterment of public action and facilitating future policy planning. However, the political and administrative appropriations of evaluations are much more complex than that in reality (Pattyn et al., 2019).

To make sense of the wide-ranging uses to which evaluation might be subjected, Vedung defined a seminal typology that is still in use today to reflect on how evaluations come into play in contemporary governance. He distinguishes five distinct types of evaluation uses. The use of evaluation is *instrumental* when evaluation results are directly implemented from a problem-solving perspective. In the *conceptual* use, evaluation comes into play gradually, enlightens its users and alters their general cognitive and normative frames to enhance their reflexive capacities around the issue. In contrast, stakeholders make use of evaluation in a *legitimizing* fashion when relying on its results to justify their pre-existing convictions. The *tactical* use of evaluations relies not on the results, but on the evaluation process itself to gain an advantage in the policy game. Examples include buying time or avoiding responsibilities by arguing that an evaluation is underway. Finally, a *discursive* use is meant to bring policy partners to the table, foster dialogue and seek conflict resolution through the evaluation (Vedung, 2000: 110-113). Below, we reflect upon the politics of evaluation. Therefore, we focus on legitimizing and tactical uses within the policy game. We use the generic term ‘strategic use’ to refer to all these possible manifestations.

The Strategic Use of Evaluation by Political Actors

There is a wide array of rationales for the political use of policy evaluations, within which we can distinguish between appropriations by actors from the executive and from the legislative branches.

For governments, an evaluation is a public assessment of their political track record. For Members of parliament, evaluations can fuel deliberations in the chambers as well as serve partisan blame games. For government and ministries, commissioning a policy evaluation might be taking the risk of a negative judgment that will be made public, so these actors do not turn their backs on evaluations as accountability instruments insofar they may also allow them to retain some control over the narrative. They can steer the process through the choice of the timing and of the dimensions to be investigated (e.g., main goals, side-effects, clients’ needs, other stakeholders’ perspectives) through an evaluation (McConnell, 2015). In so doing, governments can take control of which aspects of their office will be revealed, highlighting the achievements and downplaying the shortcomings while tampering with the policy learning potential of evaluations (Baggott, 2012). In the worst cases, evaluation commissioners can go as far as to put direct pressure on evaluators to steer the results, which shows the potential political importance of policy evaluations (Pleger et al., 2017). Other than to highlight past successes, governments might commission evaluations when taking a more defensive approach. They might do so to confirm the value of a policy when the time has come to decide on its future or when its success is publicly called into question. These kinds of processes are marked by a certain confirmation bias, and they are in line with what Carol H. Weiss identified as “the politics of program survival,” cases in which the actual evidence plays a rather minor role in the political calculus (Weiss, 1993: 103). If credit-taking is key to the political game, being attributed a policy success through an evaluation can acquire a particular meaning for politicians, as evaluations are generally executed by independent evaluators to the benefit of some scientific and professional authority. By relying on a set of specialized

procedures that sometimes achieve a high degree of technicality and theoretical and methodological sophistication, policy evaluation carries here a specific legitimacy. Evaluations serve an important outward-facing function. As Boswell (2018) has described it: evaluations can signal politicians' commitment to achieving certain goals, which can help to mobilize political support and as such support the credibility of politicians.

Compared to the executive branch, evaluations have a much weaker level of institutionalization in parliaments, generally speaking at least (Jacob et al., 2015; Stockmann et al., 2020b). This being said, a number of countries—such as Switzerland or Denmark—have indeed institutionalized ad hoc bodies as a means to exert parliamentary control over administration (Stockmann et al., 2020b). This corresponds to the classical function and the democratic ambition of evaluation as an instrument of accountability and control. In systems where such a high level of institutionalization does not prevail, it is, however, not rare for parliaments to commission evaluations and studies to assess governmental action on a specific issue, as evidenced by the multiplication of evaluations and special reports in times of crisis. The deliberative parliamentary arena is also a convenient venue for a more straightforward use of evaluations as ammunition in partisan struggles. MPs might then request evaluations to strategically oppose specific policies through, for instance, legislative committees (Bundi, 2018a). Evaluations have also been found to be used to rally political majorities in parliamentary debates. It seems that in the context of moderately contested policies, evaluations are considered as able to help settling opinions and are likely to be used for this purpose (Eberli, 2018). For the same reason, evaluation clauses are regularly introduced into laws as a means of finding political compromises, which is the case in countries as Denmark, the Netherlands or New Zealand (Jacob et al., 2015) and have proven useful in a federal consensus-style setting such as Belgium (Varone et al., 2005). In Germany, this type of mechanism has been adopted in the context of major salient policy reforms (labor market reform being an example) (Jacob et al., 2015; Speer et al., 2015). Sunset legislation with built-in evaluation requirements is another vehicle used in this regard. The idea behind this practice is that laws and government programs should periodically terminate, continuing only after an evaluation and a legislative vote to re-establish it. Implicitly, the threat of termination will function as a means to an end, the end being improved performance of an agency or program (Marvel, 1984).

In short, parliamentarians often have a plethora of evaluative information at their disposal in many different forms, including reports on managerial performance, monitoring studies, or fully fledged evaluation reports. How and when such evidence is actually used is to a large extent determined by the political agenda and ideology of the government of the day, often irrespective of the nature of the evidence, “however compelling” (Weiss et al., 2008: 33). Attributes of policy fields have been shown to strongly shape parliamentary oversight. For instance, Swiss members of parliament seem to seek more control in fields where public activities are more often delegated to non-public actors or where the need for legitimation is felt high. A field's closeness to science also has an impact in this regard (Bundi, 2018b). In addition, party dynamics matter a lot, albeit evidence is not conclusive on this issue. In some countries, asking evaluation-related questions is mainly a matter for opposition parties (e.g., in Germany), whereas in other settings (such as Flanders), coalition parties show more interest in evaluation-related information (Speer et al., 2015). Further research is needed to discern the conditions under which MPs resort to evaluation evidence in different political-administrative settings. Furthermore, there are no clear scientific results yet regarding possible distinct uses of evaluations in (varieties of) parliamentary and presidential systems. This question could also be an interesting avenue for future research.

Of course, by defining policy issues as worth being evaluated, politicians from the legislative and executive branches also contribute to setting the future policy agenda. It has been shown that the use

of evaluation in the European Parliament reflects “a clear forward-looking agenda-setting outlook rather than a backward-looking attitude” (Zwaan et al., 2016: 688). As evaluations symbolize the essence of evidence-based policy-making, they can particularly be used to move forward into the debates around emotional policy issues such as morality policies or innovative risk-taking policies. Initiating evaluation and reviews along with newly introduced controversial policies such as cannabis legalization might reassure the public and the government, thus enhancing the political feasibility of new policy paths (Hyshka, 2009).

The Administrations’ Strategic use of Evaluation

Far from constituting only a control tool of the executive and legislative branches over the administration, evaluations are also actively used strategically by administrative actors to pursue their own agendas. It is now widely recognized that public servants at all hierarchical levels enjoy a fair degree of autonomy in their implementation activities, as well as substantial agenda-setting and policy formulation power (e.g., Lavee & Cohen, 2019; Hustedt & Salomonsen, 2014 (Hustedt & Salomonsen, 2014)). Public agencies use evaluations regularly for self-assessment and monitoring to document the progress of policy implementation. However, policy evaluation can also serve the general administrative purpose of showing the public the image of a rational, sensible and responsive agency (Albæk, 1996: 21). Increasingly, extensive ‘evaluation systems’ have emerged in public sector organizations, which can be conceived as: “permanent and systematic formal and informal evaluation practices taking place and institutionalized in several interdependent organizational entities with the purpose of informing decision making and securing oversight” (Højlund, 2014a: 430). While such institutionalized evaluation provisions may give both insiders within an organization and outsiders the impression of ‘being in control’ or of ‘assurance’, the development of evaluation systems has not been free from criticism.

Scholars such as Frans L. Leeuw (2009) have pointed to the fact that evaluations are often merely conceived as an economic good, which tends to lead to ‘quick fixes’ and ‘single loop learning’ only and does not always guarantee ‘speaking truth to power’. In the same vein, scholars have criticized the short termism of evaluation systems, which entails the risk of ‘policy myopia’ and can come at the cost of asking fundamental questions about policies (Leeuw & Furubo, 2008; Raimondo & Leeuw, 2021). As it has been highlighted by institutionalists, evaluation systems first and foremost have a legitimizing role, and challenge the dominant optimistic emphasis on accountability and learning (Raimondo, 2018). In even stronger words, scholars have warned of ‘evaluation capture’, referring to situations where evaluation systems “are more beholden to the organisations they are meant to hold into account, than the public they work for” (Raimondo & Leeuw, 2021: 147). The very rules developed within and by evaluation systems can lead to dysfunctions such as goal displacement, or can ‘breed’ new interventions, which can as such actually exacerbate the very problems the evaluation systems were meant to solve (Leeuw, 2009). In light of these developments, some have called for a sceptical turn about the evaluation enterprise and the booming evaluation business (Dahler-Larsen, 2018 - also in Raimondo & Leeuw, 2021).

Their consequent autonomy also grants public agencies the necessary leeway to make strategic use of evaluations and evidence. In this regard, a wide range of tactics are available to government agencies. At the front lines of policy implementation, public agencies can decide how they use existing evaluation results (e.g., policy assessments, program evaluations). As it turns out, the instrumental use of evaluations tends to be relatively limited (Weiss et al., 2008). In addition, in organizations with a reputation of high evaluation maturity, there is a great deal of evaluation waste (e.g., Pattyn & Bouterse, 2020). Whether, to which extent and how evaluation results are used by public agencies can be conceived as a function of some crucial context conditions such as the pressure for change in the

concerned policy field, as well as the level of conflicts around the policy issue (Ledermann, 2012). The question of the effective use of evaluation results by its main recipients has been addressed thoroughly by the evaluation literature under the lead of theoretical streams on utilization-focused evaluation (Patton, 2008) and evaluation use (Alkin & King, 2017). This literature aims to provide a toolbox for the development of meaningful evaluation procedures and of practicable policy recommendations, thus attempting to maximize their social utility. In addition, public agencies do not just endure externally commissioned evaluations, but can themselves actively initiate evaluation processes to assess the current and future developments of a policy. Public agencies make extensive use of the legitimizing function of evaluations to gain intra-organizational power and influence (Widmer & Neuenschwander, 2004: 404). Just like political actors, the strategic initiation of evaluations by administrations can fulfil various objectives that tend to be overshadowed by the formal functions of evaluation as a decision-support tool. For instance, an evaluation may be useful for providing an *ex post* rationalization to unexpected policy developments. Policy implementation happens far away from the political realm, both institutionally and temporally. With time and due to field-related constraints, policy implementation sometimes moves away from the initial legislators' intent. Implementation failures may also be revealed to the media or to policymakers. In cases of an upcoming—and potentially threatening—political debate reviewing the state of implementation in a policy field, it is not rare for government agencies to order an evaluation as an opportunity to put their houses in order. This includes gathering evidence and documenting the chain of decisions that led to the existing policy path. Hence, organizations are likely to make to a justificatory use of knowledge in the face of an uncertain situation and a high-pressure context (Højlund, 2014b). From this justificatory lens, it is no surprise that organizations often strategically restrict themselves to using selective bits of evaluation findings, or reinterpret findings to fit existing preferences (Weiss et al., 2008). Accusations of policy-based evidence are neither uncommon in the evaluation field.

Finally, public servants are as likely as politicians to use evaluations to advance their policy agendas and pull the implementation in a desired direction. Given the highly tangled nature of evidence and political considerations in decision-making processes, policy evaluation constitutes an ideal tool to provide a rational packaging for specific policy scenarios (Mavrot & Sager, 2018). Policy options are here labelled as value-neutral “evidence-based” choices. In the case of international organizations, it has been shown that administrations detain even more power than member-states in influencing the evaluation process. Studies highlight the significant “resources for influence” that administrations hold, among others direct interactions with evaluators, access to evaluation drafts and considerable staff resources (Eckhard & Jankauskas, 2019). Research on international organizations thus confirms the notion that evaluation also “serves *ex ante* political interests” of different types of policy actors, and is an integral part of bureaucratic politics (Eckhard & Jankauskas, 2020: 684-685). Finally, interesting is also to refer to the so-called therapeutic function that evaluation studies can have, as shown by Lars Dorren (2021) for the case of *ex ante* analyses in infrastructural policy. Evaluations can serve as a source of validation which can bestow upon people the confidence to act. Merely conducting such studies gives participants the opportunity to reflect on their preferences, which can help them to reach a state of mind that enables them to make decisions. The fact that the process of evaluation itself can already have effects on policy stakeholders and alter their behaviour without regard to the final content of the results has also been coined as the “procedural use of evaluation” (Vedung, 1997; Widmer & Neuenschwander, 2004: 392).

Based on the literature on the politics of evaluation reviewed in this chapter, we identify three distinct types of evaluation use by political and administrative actors. The “controlling” use refers to the strategic use of evaluations by government and administrations in the context of their continuous and routine tasks of policy formulation, policy delivery and policy controlling. The “defensive” function of

evaluation is linked to specific episodes and is activated when politico-administrative actors anticipate a political debate on a particular topic and seek to strengthen their position through an evaluation. The “proactive” function of evaluation serves future agenda-setting plans of politico-administrative actors. Table 1 provides an overview of these overarching categories.

Table 1: The politics of evaluation – Political and administrative uses of evaluation

	Controlling	Defensive	Proactive
Political actors			
<i>Executive (governments, ministries)</i>	Highlighting successes	Seeking confirmation	Agenda-setting Preparing reforms
<i>Legislative (MPs, committees)</i>	Reviewing administrative activity	Rallying majorities	Agenda-setting Partisan attacks against government
Administrative actors	Documenting and monitoring policy implementation	Consolidation before political debate	Agenda-setting Pulling implementation in a direction

Conclusion: The Politicization of Evaluations

Exploring the politics of evaluation reveals how evaluations—understood as a set of established knowledge and activities aimed at formulating an evidence-based assessment on the value of public action—are strongly embedded in the political game. This chapter has provided an overview of the different possible uses of evaluation in the policy process by reviewing their controlling, defensive and proactive functions. Whether for oversight purposes, to consolidate one’s position or to prepare the ground for future political agenda, evaluations have grown to be one aspect of policy bargaining processes. Both administrators and political officials, incumbents as well as opponents, have learned to make strategic use of evaluations. In the process, evaluations can become politicized, that is, subordinated to partisan logic and political stakes (Lagroye, 2003). This is due to their particular proximity to the political system and to their focus on ongoing, sometimes disputed policy programmes that intrinsically relate to political power games. However, although evaluations are not immune to strategic uses, this does not detract from their scientific and social value. As a praxis relying on a set of established professional standards (Widmer, 2004), evaluation has tools and techniques at its disposal to distance itself from politics and that are related to theoretical, methodological and procedural soundness (L. E. Pleger & Hadorn, 2018). They can also rely on specific communication techniques to keep the interactions with politico-administrative actors under control (Sager et al., 2021). At a more macro level, in intervening in the policy pathway, evaluation is also able to induce major changes for social betterment, as many authors have theorized (e.g., Henry & Mark, 2003).

Hence, evaluation aims at providing a transparent and evidence-based analysis to improve the conceptualization, planning and implementation of public policies. It relies on scientific standards, but nevertheless remains closely entwined with the policy process because of its topics (public action), objectives (assessing current programs) and nature (applied knowledge providing a basis for future action). With the historical institutionalization of evaluation in political systems, administrative and political actors have learned to make the best of evaluations, both as a reflective instrument and as an asset in pursuing their own agenda. There lies the added value as well as the limits of policy evaluation, which provides an additional instrument of democratic oversight, to be owned and used by the existing groups and institutions in a given political system.

References

- Albæk, E. (1996). Why all this evaluation? Theoretical notes and empirical observations on the functions and growth of evaluation, with Denmark as an illustrative case. *The Canadian Journal of Program Evaluation*, 11, pp. 11–34.
- Alkin, M. C., & King, J. A. (2017). Definitions of Evaluation Use and Misuse, Evaluation Influence, and Factors Affecting Use. *American Journal of Evaluation*, 38(3), 434–450. <https://doi.org/10.1177/1098214017717015>
- Baggott, R. (2012). Policy Success and Public Health: The Case of Public Health in England. *Journal of Social Policy*, 41(2), 391–408. <https://doi.org/10.1017/S0047279411000985>
- Boswell, C. (2018). *Manufacturing political trust: Targets and performance measurement in public policy*. Cambridge University Press. <https://doi.org/10.1017/9781108367554>
- Bovens, M., & 't Hart, P. (2016). Revisiting the study of policy failures. *Journal of European Public Policy*, 23(5), 653–666. <https://doi.org/10.1080/13501763.2015.1127273>
- Bovens, M., Hart, P. 't, & Kuipers, S. (2008). The Politics of Policy Evaluation. In Goodin Robert E., Michael Moran, Rein Martin (Ed.), *The Oxford Handbook of Public Policy* (pp. 319–335). Oxford University Press.
- Bundi, P. (2018a). Parliamentarians' strategies for policy evaluations. *Evaluation and Program Planning*, 69, 130–138. <https://doi.org/10.1016/j.evalprogplan.2017.02.003>
- Bundi, P. (2018b). Varieties of accountability: How attributes of policy fields shape parliamentary oversight. *Governance*, 31(1), 163–183. <https://doi.org/10.1111/gove.12282>
- Dahler-Larsen, P. (2011). *The Evaluation Society*. Stanford University Press. <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1519293>
- Dahler-Larsen, P. (2018). The skeptical turn in evaluation. In N. Stame & J.-E. Furubo (Eds.), *Comparative policy evaluation. The evaluation enterprise: A critical view* (pp. 58–80). Routledge, Taylor & Francis Group.
- Dorren, L. (2021). *Analysis as therapy. The therapeutic function of ex ante analyses in infrastructure policy processes*. Antwerp: University of Antwerp.
- Eberli, D. (2018). Tracing the use of evaluations in legislative processes in Swiss cantonal parliaments. *Evaluation and Program Planning*, 69, 139–147. <https://doi.org/10.1016/j.evalprogplan.2016.09.007>
- Eckhard, S., & Jankauskas, V. (2019). The politics of evaluation in international organizations: A comparative study of stakeholder influence potential. *Evaluation*, 25(1), 62–79. <https://doi.org/10.1177/1356389018803967>
- Eckhard, S., & Jankauskas, V. (2020). Explaining the political use of evaluation in international organizations. *Policy Sciences*, 53(4), 667–695. <https://doi.org/10.1007/s11077-020-09402-2>
- Francesco, F. de, & Pattyn, V. (2021). Policy evaluation styles. In M. Howlett & J. Tosun (Eds.), *Routledge international handbooks. The Routledge handbook of policy styles* (pp. 408–421). New York: Routledge.
- Furubo, J.-E., & Sandahl, R. (2002). Introduction. A diffusion perspective on global developments in evaluation. In J.-E. Furubo, R. C. Rist, & R. Sandahl (Eds.), *Comparative policy analysis series. International atlas of evaluation* (pp. 1–23). Transaction Publ.
- Head, B. W. (2008). Three Lenses of Evidence-Based Policy. *Australian Journal of Public Administration*, 67(1), 1–11. <https://doi.org/10.1111/j.1467-8500.2007.00564.x>

- Henry, G. T., & Mark, M. M. (2003). Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. *American Journal of Evaluation*, 24(3), 293–314.
<https://doi.org/10.1177/109821400302400302>
- Højlund, S. (2014a). Evaluation use in evaluation systems – the case of the European Commission. *Evaluation*, 20(4), 428–446. <https://doi.org/10.1177/1356389014550562>
- Højlund, S. (2014b). Evaluation use in the organizational context – changing focus to improve theory. *Evaluation*, 20(1), 26–43. <https://doi.org/10.1177/1356389013516053>
- Hustedt, T., & Salomonsen, H. H. (2014). Ensuring political responsiveness: politicization mechanisms in ministerial bureaucracies. *International Review of Administrative Sciences*, 80(4), 746–765.
<https://doi.org/10.1177/0020852314533449>
- Hyshka, E. (2009). Turning failure into success: what does the case of Western Australia tell us about Canadian cannabis policy-making? *Policy Studies*, 30(5), 513–531.
<https://doi.org/10.1080/01442870902899962>
- Jacob, S., Speer, S., & Furubo, J.-E. (2015). The institutionalization of evaluation matters: Updating the International Atlas of Evaluation 10 years later. *Evaluation*, 21(1), 6–31.
<https://doi.org/10.1177/1356389014564248>
- Kunseler, E.-M., & Vasileiadou, E. (2016). Practising environmental policy evaluation under co-existing evaluation imaginaries. *Evaluation*, 22(4), 451–469.
<https://doi.org/10.1177/1356389016668099>
- Lagroye, J. (2003). Le processus de politisation. In J. Lagroye (Ed.), *La politisation* (pp. 359–372). Editions Belin.
- Lavee, E., & Cohen, N. (2019). How street-level bureaucrats become policy entrepreneurs: The case of urban renewal. *Governance*, 32(3), 475–492. <https://doi.org/10.1111/gove.12387>
- Ledermann, S. (2012). Exploring the Necessary Conditions for Evaluation Use in Program Change. *American Journal of Evaluation*, 33(2), 159–178. <https://doi.org/10.1177/1098214011411573>
- Leeuw, F. L. (2009). Evaluation: A Booming Business but is it Adding Value? 1. *Evaluation Journal of Australasia*, 9(1), 3–9. <https://doi.org/10.1177/1035719X0900900102>
- Leeuw, F. L., & Furubo, J.-E. (2008). Evaluation Systems: What Are They and Why Study Them? *Evaluation*, 14(2), 157–169. <https://doi.org/10.1177/1356389007087537>
- Marvel, M. K. (1984). Sunset: An Early Evaluation. *Public Choice*, 42(2), 193–196.
- Mavrot, C., & Sager, F. (2018). Vertical epistemic communities in multilevel governance. *Policy & Politics*, 46(3), 391–407. <https://doi.org/10.1332/030557316X14788733118252>
- McConnell, A. (2010). Policy Success, Policy Failure and Grey Areas In-Between. *Journal of Public Policy*, 30(3), 345–362. <https://doi.org/10.1017/S0143814X10000152>
- McConnell, A. (2015). What is policy failure? A primer to help navigate the maze. *Public Policy and Administration*, 30(3-4), 221–242. <https://doi.org/10.1177/0952076714565416>
- Palumbo, D. J., & Hallett, M. A. (1993). Conflict versus consensus models in policy evaluation and implementation. *Evaluation and Program Planning*, 16(1), 11–23. [https://doi.org/10.1016/0149-7189\(93\)90033-5](https://doi.org/10.1016/0149-7189(93)90033-5)
- Patton, M. Q. (2008). *Utilization-focused evaluation* (4. ed.). SAGE.
- Pattyn, V., & Bouterse, M. (2020). Explaining use and non-use of policy evaluations in a mature evaluation setting. *Humanities and Social Sciences Communications*, 7(1).
<https://doi.org/10.1057/s41599-020-00575-y>

- Pattyn, V., Peuter, B. de, & Brans, M. (2019). Why do Ministers Ask for Policy Evaluation Studies? The Case of the Flemish Government. *Politische Vierteljahresschrift*, 60(4), 701–717. <https://doi.org/10.1007/s11615-019-00211-8>
- Peters, B. G., Capano, G., Howlett, M., Mukherjee, I., Chou, M.-H., & Ravinet, P. (2018). *Designing for Policy Effectiveness*. Cambridge University Press. <https://doi.org/10.1017/9781108555081>
- Pleger, L., Sager, F., Morris, M., Meyer, W., & Stockmann, R. (2017). Are Some Countries More Prone to Pressure Evaluators Than Others? Comparing Findings From the United States, United Kingdom, Germany, and Switzerland. *American Journal of Evaluation*, 38(3), 315–328. <https://doi.org/10.1177/1098214016662907>
- Pleger, L. E., & Hadorn, S. (2018). The big bad wolf's view: The evaluation clients' perspectives on independence of evaluations. *Evaluation*, 24(4), 456–474. <https://doi.org/10.1177/1356389018796004>
- Raimondo, E. (2018). The power and dysfunctions of evaluation systems in international organizations. *Evaluation*, 24(1), 26–41. <https://doi.org/10.1177/1356389017749068>
- Raimondo, E., & Leeuw, F. L. (2021). Evaluation systems and bureaucratic capture. Locked in the system and potential avenues for change. In B. Perrin & T. Tyrrell (Eds.), *Comparative policy evaluation: Vol 29. Changing bureaucracies: Adapting to uncertainty, and how evaluation can help*. Routledge.
- Sager, F. (2017). Policy evaluation and democracy: Do they fit? *Evaluation and Program Planning*. Advance online publication. <https://doi.org/10.1016/j.evalprogplan.2017.08.005>
- Sager, F., Hadorn, S., Balthasar, A., & Mavrot, C. (2021). *Politikevaluation. Eine Einführung*. Springer.
- Sager, F., & Mavrot, C. (2021). Participatory vs expert evaluation styles. In M. Howlett & J. Tosun (Eds.), *Routledge international handbooks. The Routledge handbook of policy styles* (pp. 395–407). Routledge.
- Scriven, M. (1991). *Evaluation thesaurus* (Hesaurus. Bla). Blackwell's, Oxford.
- Speer, S., Pattyn, V., & Peuter, B. de (2015). The growing role of evaluation in parliaments: holding governments accountable? *International Review of Administrative Sciences*, 81(1), 37–57. <https://doi.org/10.1177/0020852314546583>
- Stame, N. (2003). Evaluation and the Policy Context: The European Experience. *Evaluation Journal of Australasia*, 3(2), 36–43. <https://doi.org/10.1177/1035719X0300300210>
- Stame, N. (2012). Evaluation and learning in the current crisis. *Evaluation Connections*, January, 3–4.
- Stephenson, P. J., Schoenefeld, J. J., & Leeuw, F. L. (2019). The Politicisation of Evaluation: Constructing and Contesting EU Policy Performance. *Politische Vierteljahresschrift*, 60(4), 663–679. <https://doi.org/10.1007/s11615-019-00212-7>
- Stockmann, R., Meyer, W., & Taube, L. (2020a). *The Institutionalisation of Evaluation in Europe* (1. ed. 2020). Springer International Publishing; Imprint: Palgrave Macmillan.
- Stockmann, R., Meyer, W., & Taube, L. (2020b). The Institutionalisation of Evaluation in Europe: A Synthesis. In R. Stockmann, W. Meyer, & L. Taube (Eds.), *The Institutionalisation of Evaluation in Europe* (pp. 483–522). Springer International Publishing. https://doi.org/10.1007/978-3-030-32284-7_19
- Varone, F., Jacob, S., & Winter, L. de (2005). Polity, Politics and Policy Evaluation in Belgium. *Evaluation*, 11(3), 253–273. <https://doi.org/10.1177/1356389005058475>
- Vedung, E. (1997). *Public policy and program evaluation*. New Brunswick, NJ: Transaction.

- Vedung, E. (2000). Evaluation Research and Fundamental Research. In R. Stockmann (Ed.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder* (pp. 103–126). Leske + Budrich.
- Vedung, E. (2010). Four Waves of Evaluation Diffusion. *Evaluation*, 16(3), 263–277.
<https://doi.org/10.1177/1356389010372452>
- Weaver, L., & Cousin, B. (2004). Unpacking the Participatory Process. *Journal of MultiDisciplinary Evaluation*, 1(1), 19–40.
- Weiss, C. H. (1993). Where Politics and Evaluation Research Meet. *Evaluation Practice*, 14(1), 93–106.
<https://doi.org/10.1177/109821409301400119>
- Weiss, C. H., Murphy-Graham, E., Petrosino, A., & Gandhi, A. G. (2008). The Fairy Godmother—and Her Warts. *American Journal of Evaluation*, 29(1), 29–47.
<https://doi.org/10.1177/1098214007313742>
- Widmer, T. (2004). The development and status of evaluation standards in western Europe. *New Directions for Evaluation*, 2004(104), 31–42. <https://doi.org/10.1002/ev.134>
- Widmer, T., & Neuenschwander, P. (2004). Embedding Evaluation in the Swiss Federal Administration. *Evaluation*, 10(4), 388–409. <https://doi.org/10.1177/1356389004050283>
- Zwaan, P., van Voorst, S., & Mastenbroek, E. (2016). Ex post legislative evaluation in the European Union: questioning the usage of evaluations as instruments for accountability. *International Review of Administrative Sciences*, 82(4), 674–693. <https://doi.org/10.1177/0020852315598389>