The Hidden Politics of Evaluation:
Towards a Smarter State?

By:

Perla Carolina Gris Legorreta

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

The University of Sheffield
Faculty of Social Sciences
Department of Politics

September 2015

Abstract

The neutrality of evaluation processes has been widely discussed in both the academic and practitioner literature. The abstraction of assessment mechanisms from the reality of public policy arguably seems naïve and disconnected from the actual role of evaluation. It is in exactly this context that this thesis explores the politics of evaluation in terms of both an internal process and an external procedure within accountability frameworks. By focusing on the political nature of evaluation in healthcare in Mexico and the United Kingdom this thesis sheds new light on the complex relationships and inter-dependencies that dominate the architecture of modern governance. The core finding of the research presented in this thesis is that the hidden politics of evaluation matters because of the way in which the institutional framework provides actors with spaces of discretion that allow them to influence the process and outcomes of evaluation. This finding, and the research that underpins it, adds to our understanding in terms of the potential gap between what might make 'good policy', as opposed to what might make 'good politics.'

Acknowledgements

Contents

List of Tables, Figures and Boxes

List of Abbreviations

AEA	American Evaluation Association
AEP	Annual Evaluation Programme
C&GA	Comptroller and Auditor General
CDF	Cancer Drugs Fund
CONEVAL	National Council for the Evaluation of Social Development Policy
CQC	Care Quality Commission
CRS	Cancer Reform Strategy
CRUK	Cancer Research UK
DfID	Department for International Development
DOH	Department of Health
EEF	Education Endowment Foundation
FPA	Federal Public Administration
GP	General Practitioner
HTA	Health Technology Assessment Programme
IADB	Inter-American Development Bank
ICAI	Independent Commission for Aid Impact
INEE	National Institute for the Evaluation of Education
INSP	National Institute of Public Health
IOSC	Improving Outcomes: A Strategy for Cancer
LICONSA	Social Milk Supply Programme
M&E	Monitoring and Evaluation
MMU	Medical Mobile Units
MP	Member of Parliament
NAO	National Audit Office
NHS	National Health Service
NCIN	National Cancer Intelligence Network
NCEI	National Cancer Equality Initiative
NICE	National Institute for Clinical and Health Excellence
NGO	Non-governmental Organisation
NPM	New Public Management
OPA	Output and Performance Analysis
OECD	Organisation for Economic Co-operation and Development
ORP	Operational Rules of the Programme
PAC	Public Accounts Committee
PAN	National Action Party
PES	Performance Evaluation System
PFI	Private Finance Initiative
PHE	Public Health England
PM	Prime Minister
PRD	Party of the Democratic Revolution
PRI	Institutional Revolutionary Party
PSA	Public Service Agreements
QDS	Quarterly Data Summary
RCT	Randomised Controlled Trial

ROAMEF	Rationale, objectives, appraisal, monitoring, evaluation, and feedback
SAO	Supreme Audit Office
SHA	Strategic Health Authority
SICALIDAD	Health Quality System Programme
SMART	Specific, measurable, achievable, relevant, and timed
SPE	Specific Performance Evaluation
TOR	Terms of Reference
UKPAC	UK Public Affairs Council
VFM	Value for Money
WB	World Bank

Chapter 1. Introduction

The evolving nature of the modern state has encouraged the promotion of internal and external mechanisms of control through which the general public, as well as politicians, officials, and interest groups, can evaluate the capacity of the state to respond to social demands. In the context of modern governance, the use of evaluation techniques has arisen as a common instrument to identify the effectiveness of public policy. However, the politics[1] involved in evaluation requires being analysed in depth in the sense of grasping how various actors involved might influence the evaluation process and its outcomes to promote their interests, to seek political advantage, or to produce changes within the structure of the state.[2]

The 'audit explosion' described by Power and the adoption of tools such as monitoring and evaluation are only some examples of the way in which managerial models can modify how policy is conceptualised, implemented, and assessed.[3] Regarding the latter, it is useful to reflect upon the political values that drive evaluation.[4] The definition of 'success' and the criteria used for its identification in public policy are ideological constructions resulting from negotiations between various actors.

The purpose of this thesis is to expose the hidden politics of evaluation and therefore the relationship between evaluation on the one hand and politics on the other, by looking at the evaluation policies implemented in two different countries: the United Kingdom (UK) and Mexico. Through four case studies the

---

[1] This thesis adopts the concept of politics proposed by Lasswell *(who gets what, when and how)* understood as the way in which stakeholders of evaluation pursuit their interests under the conditions imposed by the institutional framework. A more detailed development of the concept of politics can be found in Chapter 2, section 2.3.

[2] Actors involved in evaluation can include 'the commissioners or funders (who may set the parameters for the work and the expectations for its outcome), the key audiences for the work, those involved in advisory panels (who may influence the quality standards applied), and participants in the evaluation (who may bring a range of knowledge and views to bear, as well as influence how easily it can be conducted.' Lonsdale, J. 2008. "Balancing Independence and Responsiveness: A Practitioner Perspective on the Relationships Shaping Performance Audit." *Evaluation* 14(2):227-48. P. 228.

[3] Power, M. (1997). *The Audit Society: Rituals of Verification*. Oxford, UK: Oxford University Press.

[4] Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106.

role of those actors involved in the use or deployment of evaluation processes and findings is analysed. The focus on the political nature of evaluation in healthcare in these countries allows the bringing of new light on the complex relationships and interdependencies that dominate the architecture of modern governance. It adds to our understanding in terms of the potential gap between what might make 'good policy', as opposed to what might make 'good politics'. As it is discussed throughout this thesis, the political nature of evaluation has been acknowledged by the scholars in the field.[5] Evaluation entails the adoption of political values that affect its processes and outcomes. However, in the practical domain this political nature has been hiding behind the technocratic label given to evaluation, seen as an objective and a rational instrument to analyse different features of public policy.[6] Therefore, it seems relevant to uncover the politics imbued in the process of defining the success or failure of a policy, as well as in the use, dismissal, or promotion of the outputs of this activity. Politics in the context of this thesis refers to the way in which actors use their resources and power to influence how evaluation is planned, conducted, and disseminated.

Evaluation as a risk-reduction mechanism and as an aid for accountability processes can be located both *within* and *outside* the government sphere. Internally, evaluation can serve the purpose of verifying and controlling the performance of organisations and individuals. Externally, it can provide legislative bodies and civil society (predominantly) with an instrument for holding the government to account and scrutinising the effectiveness of public policy.

Due to the centrifugal pressures of New Public Management (NPM) reforms, governments around the world require more sophisticated tools to obtain reliable and objective information about their actions. Evaluation provides a

---

[5] Banner, D. K. 1974. "The politics of evaluation research." *Omega* 2(6):763-74, Palumbo, D. J. 1987. *The Politics of Program Evaluation*. California, US: SAGE, Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106.

[6] This idea relates to the evidence-based policymaking approach, which 'indicates the continuing force of optimism about the potential to achieve social progress through the application of reason.' Sanderson, I. 2002. "Evaluation, Policy Learning and Evidence-Based Policy Making." *Public Administration* 80(1):1-22. P. 19.

centripetal pressure that resonates with the idea of 'filling in' the 'hollowing out' and yet there is very little known about evaluation processes and even less about the politics of evaluation.[7] This idea of the 'hollowed-out state' in the context of evaluation is relevant because the information derived from this activity intends to provide technical knowledge for improving control and diminishing risk.[8] The potential application of evaluation tools demands a more detailed analysis in terms of politics because its outputs can be perceived by stakeholders as either an aid or a threat to their agendas. Therefore, the implementation of evaluation systems may imply the establishment of adversarial relationships among stakeholders as a result of the conditions needed for the functioning of these systems, i.e. cooperation, negotiation, resources, or control.

This thesis is based on the argument that the institutional framework established for evaluation limits the way in which politics influences the process. The rules and procedures underpinning evaluation can predict and to some extent shape the behaviour of the actors involved. The study of two different scenarios reveals how different institutional arrangements may produce a different configuration of power, e.g. by allowing an actor a privileged position in the control of the process.

The core finding of this research is that the hidden politics of evaluation matters because of the way in which the institutional framework provides actors with spaces of discretion that allow them to influence the process and outcomes of evaluation. In the UK, for example, the institutional framework provides government ministries with an important amount of discretion for the evaluation of their policies and it favours the participation of

---

[7] 'Hollowing-out' refers to the idea 'that government is losing its ability to govern because of increasing domination by autonomous, self- organising policy networks.' Taylor, A. (2000). Hollowing out or filling in? Taskforces and the management of cross-cutting issues in British government. *The British Journal of Politics and International Relations, 2*(1), 46-71. P. 46.
[8] 'Filling in' the 'hollowing out' of the state through policy evaluation can be linked to the idea that 'governments or other public organisations, like local authorities, have at their disposal a toolbox that allows them to deploy different instruments in various combinations for the circumstances they face, the state or public authority gives those in office a set of facilities other actors do not have.' John, P. (2011). *Making Policy Work*. New York, U.S.:Routledge. P. 7.

external actors in the process.[9] In this case, it is also possible to identify a strong audit system with which work is perceived as legitimate and highly influential in the debate about the effectiveness of government.[10] In contrast, the Mexican case is characterised by an excessively centralised evaluation policy, with which the Federal Government exercises control over core elements of the process. This seems to interfere with the objectivity and independence of this activity, particularly in a context in which evaluation has been given a legitimising function. The research findings suggest that although the institutional framework can determine rules and procedures for the interaction of stakeholders, discretionary spaces found in this framework are used for pursuing particular agendas. The implication of this core finding is that it is not only impossible to depoliticise evaluation, but it is also undesirable. As is argued in this thesis, politics may be either a positive or negative force for evaluation that needs to be acknowledged, understood, and integrated into its rationale in a more open manner; the hidden politics of evaluation needs to be brought into the open to foster its utility as a governance tool.

1.1 What? The Focus of this Research

Public policy can be understood as 'a statement by government – at whatever level – of what it intends to do about a public problem.'[11] It entails the use and allocation of public resources and the development of different actions to address problems that affect society.[12] In most cases, these decisions will benefit a large section of society, while other sectors of the population might be neglected due to the intended or unintended effects of a policy.[13] As a result, the analysis of public policy has emerged as a sub-discipline of political science that studies the different activities undertaken by governments.

[9] The *Evaluation of Nurse and Pharmacist Independent Prescribing*, for example, comprised the development of a multi-stakeholder workshop 'to consider and prioritise the preliminary study findings and implications', while it also involved a survey to 358 pharmacist independent prescribers. See: Department of Health. 2011a. "Evaluation of nurse and pharmacist independent prescribing." London, UK.

[10] Talbot, C., and J. Wiggan. 2010. "The public value of the National Audit Office." *International Journal of Public Sector Management* 23(1):54-70.

[11] Birkland, T. A. 2011. *An Introduction to the Policy Process: Theories, Concepts, and Models of Public Policy Making*. New York, US: M.E. Sharpe. P. 9.

[12] Miyakawa, T. 1999. *The Science of Public Policy*. London, UK: Routledge.

[13] See: Weimer, D. L., and Vining, A. R. (2004). *Policy Analysis: Concepts and Practice* (4th ed.). New Jersey, U.S.: Pearson Prentice Hall.

As one of the stages of the policy process, evaluation aims to assess different aspects of public policy, such as its impact, its cost-effectiveness, and its results. As Picciotto states, 'evaluation determines the merit, worth and value of things [...] it consists in collecting relevant evidence, identifying suitable evaluative standards and using methods of analysis that are valid and fair.'[14] It represents a tool of governance that allows stakeholders to cope with the political pressures found in the public sector, both in terms of policy improvement as well as for accountability purposes.[15] Evaluation can provide elements to oversee and control public organisations, to promote coordination within the public sector, and to establish clear criteria for defining the success or failure of a policy.

This PhD thesis is located within the field of policy analysis and, more specifically, within evaluation studies. Its main purpose is to contribute to the evaluation studies literature by addressing the relationship between evaluation and its political context. Reflecting on this topic allows understanding the potential uses that actors can make of evaluation to pursue their interests. Politicians, for example, can find in evaluation a valuable asset for them 'to claim credit and avoid blame from voters.'[16] Nonetheless, the analysis of the hidden politics of evaluation is not constrained to the political use of findings. As discussed in this thesis, the political influence that actors can exercise over the evaluation process will depend on their agenda, the resources at their disposal, and the rules established for the performance of this activity.

The political nature of evaluation needs to be studied in more detail for different reasons. The conditions in which evaluations are designed and implemented are underpinned by the values defended by the actors that control totally or partially the process, but *what does the hidden politics of evaluation imply?* In the first place, politics demands looking at the way in which power is disseminated among actors involved in evaluation. This refers, for example, to how politicians,

---

[14] Picciotto, R. (2005). Policy Coherence and Development Evaluation: Issues and Possible Approaches *Fostering Development in a Global Economy. A Whole of Government Perspective*. Paris, France: Organisation for Economic Co-operation and Development. P. 4.
[15] Peters, B. G. 2015. *Advanced Introduction to Public Policy*. Cheltenham, UK: Edward Elgar Publishing.
[16] Hood, C. 2002. "The Risk Game and the Blame Game." *Government and Opposition* 37(1):15-37. P. 17.

operators, and evaluators exercise their power to shape the way in which policy is assessed. This allocation of power can be perceived in how resources are applied and how rules and procedures are established. Undeniably, evaluation is underpinned by a political logic, because as Weiss stated, policies are political constructions.[17] Nonetheless, in practice the politicisation of evaluation is frequently disguised behind technocratic, rational, and objective dimensions associated with it. As discussed in Chapter 2, the recognition of the political nature of evaluation does not seem to be acknowledged (at least not explicitly) in the establishment of arrangements and conditions for the performance of this activity, even though political relationships influence the conduction of this activity.

Another important issue to highlight are the conditions that the institutional framework establishes for actors to exercise their power. Not only the rules, procedures, and institutions set for evaluation are the results of political arrangements, but also the outcomes of these agreements will also define the extent to which actors can participate, i.e. selecting programmes to evaluate. From this perspective, the understanding of the relationship between evaluation and politics can be summarised in two major points:

- Politics can be understood as the way in which actors involved in evaluation can exercise their power and resources to shape the process according to their interests.
- Despite the acknowledgement of the political component of evaluation, in practice there is little recognition of how the set of rules established for evaluation is affected by the political interests of the actors involved in this process.

Evaluation might affect the interests of multiple stakeholders. Consequently, the allocation of power for deciding how evaluation is designed and implemented can be subject to political struggles for the prioritisation of those issues that are particularly relevant for a group, i.e. efficiency. Moreover, evaluation outputs can influence key political decisions such as the allocation of funding, the continuity of a programme, or the designation of a public officer. This gives evaluation an important leverage in the public sector, and a positive or negative evaluation can therefore become a weapon for political purposes.

---

[17] Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106.

This thesis discusses the relationship between the political nature of evaluation and the institutional framework. This is understood to be the set of rules defined for the performance of evaluation that influence how power and resources are allocated, as well as the interaction among stakeholders. The actors that have access to the definition of these rules have a privileged position in relation to others. For instance, evaluation commissioners have the prerogative to define how an evaluation will be conducted and the methods that will be used.

The major purpose of this research is to generate new knowledge to fill the gap between evaluation on the one hand, and politics on the other by addressing three main elements:
- The areas of the evaluation process susceptible to political influence
- The manifestation of the political nature of evaluation in the empirical context
- The effect of the institutional framework in the political relationships that are established as a consequence of evaluation

Although the focus of this thesis is on evaluation, auditing is also considered a relevant object of study because of its role as a control/scrutiny mechanism, particularly at present when its conceptualisation has transcended the financial dimension and has become closer to evaluation.[18] The theoretical differences between 'audit' and 'evaluation' are more perceivable in terms of scope, methods, and rationale.[19] According to Leeuw these distinctions are related to three major elements: the variables investigated, the methods and techniques used and the underlying concept of these activities.[20] Regarding the variables investigated, auditing 'emphasizes the process of checking or verifying records to find out whether they are consistent with agreed-upon standards.'[21] Evaluations instead focus on 'goal achievement and the intended and unintended side effects of policies and programs.'[22]

---

[18] Pollitt, C., and H. Summa. 1996. "Performance Audit and Evaluation: Similar Tools, Different Relationships?" *New Directions for Evaluation* 1996(71):29-50.
[19]. See: Chelimsky, E. Ibid."Auditing and Evaluation: Whither the Relationship?":61-67.
[20] Leeuw, F. L. Ibid."Auditing and Evaluation: Bridging a Gap, Worlds to Meet?":51-60.
[21] Brooks, R. A. Ibid."Blending two cultures: State legislative auditing and evaluation."15-28. P. 16.
[22] Leeuw, F. L. Ibid."Auditing and Evaluation: Bridging a Gap, Worlds to Meet?":51-60. P. 51.

The methods used in auditing are predominantly financial, such as cost-benefit analysis or value for money (VFM) studies[23], while evaluation has been supported by the use of social research methods for its development.[24] Finally, it might be in the rationale of both activities when the most notorious differences arise:

> Auditing, conceived as an accounting instrument, focuses fundamentally on the processes and on the verification of a series of criteria within an established analytical framework, highlighting the deviations found and providing information about the non-compliance of the norms and the relationship between activities and resources. Evaluation, on the other hand, has focused its attention on the results of programmes and policies in the aim of generating evidence about their effectiveness.[25]

Making this distinction clear at this point is necessary because the findings derived from these activities are useful for identifying the pertinence of the decisions and actions taken in the context of policymaking. Moreover, both the process and the outputs of evaluation and auditing are susceptible to political influence. Looking at how these instruments interact in the practical domain is important because its integration or dissociation will impact on the overall effect of control mechanisms in the public sector and on its vulnerability to political influence.

In the context of the 21st Century, the study of the politics of evaluation is crucial for observing the extent to which the state has become smarter by not only taking advantage of the managerial and political benefits that both audit and evaluation offer, but also by being able to address the potential challenges that politics represents for the adoption of these instruments.[26]

---

[23] Chelimsky, E. 1985. "Comparing and Contrasting Auditing and Evaluation: Some Notes on their Relationship." *Evaluation Review* 9(4):483-503.

[24] Shaw, I., J. C. Greene, and M. M. Mark. 2006. *Handbook of Evaluation: Policies, Programs and Practices.* London, UK: SAGE.

[25] Gris Legorreta, P. C. 2013. "Auditoría y Evaluación: ¿Hacia un tercer enfoque?" *XIII Certamen Nacional de Ensayo Sobre Fiscalización y Rendición de Cuentas.* P. 56.

[26] The use of evaluation and performance monitoring information in the public sector has been an important initiative of international organisations such as the WB, the IADB and the OECD. See Curristine, T., Z. Lonti, and I. Joumard. 2007. "Improving Public Sector Efficiency: Challenges and Opportunities " *OECD Journal on Budgeting* 7(1):1-42.

1.1.1 Research Questions and Hypothesis

The core question at the heart of this thesis is therefore: *What is the relationship between the political nature of evaluation and its institutional framework?*

The hypothesis proposed in this thesis is that the institutional framework can affect the degree of political influence in policy evaluation and the consequences for the utility and legitimacy of this activity. Decisions and actions taken in the context of the evaluation are subject to the political influence of those key actors participating in this process. They are not only involved in the development of evaluations and the use and dissemination of findings, but also in the establishment of rules and procedures that shape their behaviour in the context of this activity. The relevance of the core theme of this thesis is that it allows understanding how different institutional arrangements for evaluation lead to different outcomes, e.g. how rules defined for the commissioning of evaluations affect the level of discretion that evaluators have for their task. The nuances of this hypothesis are addressed in detail by looking at how health policy is evaluated in two different contexts.

The political arrangements that take place during evaluation may be promoted or limited by the boundaries established by the institutional framework. This concerns its potential as input for the political discussion and for the construction of arguments. Changes in the way in which public organisations deliver services can also be shaped by the evidence provided by evaluation. It can be conceived as a tool of governance to support the position of stakeholders in a context in which political values and beliefs are constantly in conflict. While evidence 'helps policymakers make better decisions, and achieve better outcomes...'[27] the use of evaluation has not always derived a successful outcome because there are organisational, political, and managerial variables that might have influence.[28] For this reason, understanding the role of politics in evaluation can generate new knowledge about why sometimes evaluation findings are not

---

[27] Davies, P. 2012. "The State of Evidence-Based Policy Evaluation and its Role in Policy Formation." *National Institute Economic Review* 219(1):R41-R52. P. R42.

[28] Jones, H. 2012. "Promoting evidence-based decision-making in development agencies." Overseas Development Institute.

reflected as better policies or greater accountability.[29]

The research question represents a general guideline for the development of this thesis; nonetheless, to reach a more specific level of analysis it is necessary to disaggregate it into more operational elements (Table 1.1). These questions will be addressed through a thematic framework proposed for the analysis of the politics of evaluation (see Chapter 2, Section 2.3.2). The development of secondary questions highlights the relevance of the institutional framework in relation to the political nature of evaluation processes. The linkage between the secondary questions is based on the identification of key elements of the evaluation process that can bring new light to the study of the political nature of evaluation. This refers, for example, to the characteristics of evaluation instruments, the resources available for evaluation, the utility perceived by stakeholders, and the conflicts that arise between the administrative and political dimensions of evaluation.

---

[29] It is possible to find in the literature several critiques to the evidence-based policymaking approach in relation to its incapacity to promote better outcomes in government. See for example: Greenhalgh, T., and J. Russell. 2009. "Evidence-Based Policymaking: A Critique." *Perspectives in Biology and Medicine* 52(2):304-18, Marston, G., and R. Watts. 2003. "Tampering With the Evidence: A Critical Appraisal of Evidence-Based Policy-Making." *The Drawing Board: An Australian Review of Public Affairs* 3(3):143-63.

Table 1.1 Secondary Research Questions to Explore the Hidden Politics of Evaluation

Key Variables	Six Secondary Research Questions	Relevance	Section of the thesis in which the question will addressed
Institutions	1. Are there efficient institutional frameworks that assure an effective assessment of public policy?	The institutional framework in which an evaluation policy is implemented defines not only the general premises that need to be followed, but also the expected use of the evaluation results and the actors responsible and accountable for this activity.	See sections 4.1 and 5.1
Values	2. Which values are promoted by the different stakeholders involved in evaluation? How do these values interact?	As a political activity, evaluation is underpinned by different values (e.g. economy, efficiency), it seems relevant to identify them and understand how these interact in an arena where multiple stakeholders intervene and values might be in conflict.	See sections 4.2 - 4.3 and 5.2-5.3
Outputs	3. Does the evaluation policy, in both countries produce (in a systematic way) relevant and convenient information that can be used for the policymaking process?	One way to identify the political influence of evaluation is to analyse the kind of results it generates. By studying and comparing the outputs of evaluation it will be possible to understand if it produces objective information or if there can be identified biased results that might favour the interests of stakeholders.	See sections 6.1 and 6.2
Outcomes	4. When the outputs of evaluation are used, is there any effect on the improvement of policymaking?	The ultimate goal of evaluation is to produce knowledge that can be used in the policymaking process; however, the political interests of the stakeholders of evaluation might determine its use or misuse.	See section 6.3

Table 1.1 *Continued*

Key Variables	Six Secondary Research Questions	Relevance	Section of the thesis in which the question will addressed
Culture	5. Which is the role of policy evaluation in order to consider it useful for the policymaking process?	After analysing the different areas in which it can be identified the influence of politics in evaluation, it is relevant to discuss the role of evaluation in a context of political pressures and of opposed interests, particularly with the purpose of formulating recommendations for strengthening the policymaking process.	See section 7.3
Politics	6. In terms of 'good policy' and 'good politics', which of these issues predominate in the evaluation agenda of both countries?	The political influence in policy evaluation has an effect on the use of the evaluation findings, therefore, it is important to identify in which extent this influence has conditioned the utility of evaluation in both case studies.	See section 7.1

1.1.2 Structure of the Thesis

To address the research questions of this thesis it is important to describe its structure and the general content of the seven chapters (Table 1.1). The first chapter is an introduction in which the main topics of the thesis are presented; the scope and the core questions that will guide the development of the thesis are established. The second chapter is dedicated to the theoretical framework. It presents the state of the art in relation to the analysis of governance, public policy, and evaluation studies. This framework allows the identifying of the most relevant discussions found in the literature for the study of the politics of evaluation. The transformations that have occurred within the structure of the state have fostered the development of new control mechanisms in which evaluation processes play an important role. Consequently, it is important to understand how the political nature of evaluation relates to the new processes of governance, and more importantly, how the *politicisation* of evaluation affects its role as a mechanism of control in policymaking and accountability processes.

The third chapter establishes the methodological framework of the research. It describes the tools implemented, the justification of the cases selected, and the possible limitations that need to be considered when looking at the findings. The methodology applied in this research is particularly relevant because of the intangibility of politics. The identification of key variables to study this phenomenon allows a more structured grasping of the areas of the evaluation process by which politics can influence the way stakeholders can shape this process. The fourth chapter focuses on the description of the UK evaluation system to make explicit the legal and organisational framework in which this activity occurs. It focuses on the process for evaluating and auditing public policy, programmes, and organisations. This chapter also describes the case studies selected for the UK: the Cancer Strategy and the Care-Quality Commission (CQC). The fifth chapter replicates this description for the Mexican case and the selected case studies: the Integral Quality Health System Programme (SICALIDAD) and the Health Caravans Programme.

The sixth chapter presents the analysis of the four case studies guided by the thematic framework outlined in Chapter 2 for uncovering the hidden politics of

evaluation. This chapter compares and contrasts the evaluation process in both countries, taking into consideration the rules, i.e. governance, established for this purpose.

The final chapter is dedicated to the presentation of the conclusions and the formulation of policy recommendations. The latter concerns fundamentally the strategies and mechanisms to strengthen the linkage between evaluation and policymaking, formulated on the basis of the empirical findings produced by this thesis. These recommendations identify the different roles that the stakeholders of evaluation should play to build a solid system that produces reliable feedback for policymaking. It presents the conclusions of both the theoretical and empirical analysis and it outlines the parameters of a future research agenda.

1.2 Why? The Purpose and Claims for Distinctiveness

The importance of a PhD thesis relies on its capacity to expand the existing knowledge in a particular field through the generation of new information (theoretical and empirical) that can contribute to the comprehension of different social phenomena. Considering that the study of the hidden politics of evaluation is an area of the evaluation studies field that has been little explored from an empirical approach, in contrast to other topics in the field, this thesis is pertinent in terms of analysing its implications for the theoretical and practical contexts. Although since the late seventies different scholars have acknowledged the political nature of evaluation, research about this topic has been mostly limited to the analysis of the political pressures that evaluators face and the way they respond to these.[30] The overall impact of evaluation as a mechanism for political legitimation and control of the policymaking process has not received much attention.

Two elements are worth highlighting in terms of the originality and distinctiveness of this thesis. The first one is that it conceptualises evaluation as

---

[30] See: Hedrick, T. E. 1988. "The Interaction of Politics and Evaluation." *American Journal of Evaluation* 9(3):5-14, Palumbo, D. J. 1987. *The Politics of Program Evaluation*. California, US: SAGE, Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106.

a tool for governance that allows stakeholders to cope with the increasing demands for accountability, transparency, and effectiveness in a context characterised by a fragmented bureaucracy, the delegation of a government's power, and a more intensive participation of external actors. For instance, the rising interest of governments in evaluating their policies has represented an important source of business for universities, think tanks and consultants.[31] Programme operators are interested in the identification and recognition of the goals achieved and 'in building long-term support for the program.'[32] Politicians find in evaluation evidence for supporting their decisions.[33] This evinces the multiple uses that evaluation can have according to the demands and requirements of the actors involved.[34]

The second issue is that this thesis adopts a more holistic approach to the political nature of evaluation by looking at its process and the interaction with the institutional framework. This provides two important benefits. On the one hand, it transcends the analysis of the hidden politics of evaluation based on the perception or experience of particular stakeholders, e.g. evaluators. Instead, it aims to present a more integrated analysis of how key elements of evaluation can be politically influenced by different stakeholders. On the other hand, it reveals how the establishment of rules and procedures for evaluation can shape the behaviour of stakeholders. This last point is relevant because the level of political influence in the process may have implications for the utility and legitimacy of evaluation.

1.2.1 Motivations of the Research

The drivers for the development of this research can be synthesised in two main motivations. The first one regards my professional development, primarily focused on the public sector. As a former official in the Mexican evaluation

---

[31] See: PWC. 2012. "The Business of Evidence. A report prepared for the Market Research Society." London, UK.
[32] Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106.
[33] Sanderson, I. 2002. "Evaluation, Policy Learning and Evidence-Based Policy Making." *Public Administration* 80(1):1-22.
[34] Feinstein, O. N. 2002. "Use of Evaluations and the Evaluation of their Use." *Evaluation* 8(4):433-39.

agency, I identified the potential of evaluation processes to foster change in public organisations. At the same time, this experience also allowed me to recognise how the acknowledgement and management of the political factor can maximise the benefits of evaluation. The second motivation relates to the academic perspective. As stated, there are substantial gaps in the knowledge about the political nature of evaluation. Consequently, it represents an important motivation to contribute to the evaluation studies field by providing empirical evidence about the relationship between the politics of evaluation and the institutional framework. This research recognises the need for studying this phenomenon from an integrated perspective and not merely from the point of view of practitioners and evaluators, which although is clearly valuable is insufficient for its holistic comprehension.

This research also aims to feed the discussion with information about the role of evaluation as a control and risk-reduction mechanism. Undeniably, political institutions around the world are suffering a credibility crisis that has provoked a serious disengagement between citizens and political institutions.[35] This is perceivable in the two countries on which this thesis focuses. According to a recent survey, in Mexico 65% of the interviewees declared having a small interest in politics.[36] The UK presents a similar situation. The latest British Social Attitudes survey shows that 64% of the people considered themselves to be having either some, not much, or no interest at all in politics.[37] This suggests that governments need to develop mechanisms to rebuild trust in society and that evaluation can play a useful role in the legitimisation of public action. The selection of these two contexts provides new insights into how (and why) evaluation mechanisms have developed under different institutional arrangements and the results that these have produced in the context of governance.

---

[35] This represents 'a major governance problem, reflecting a decline in the legitimacy of government and, thus, problems in developing and implementing policy.' Fawcett, P., and D. Marsh. 2014. "Depoliticisation, governance and political participation." *Policy & Politics* 42(2):171-88. P. 171.
[36] Secretariat of the Interior. 2012. "Fifth National Survey about Political Culture and Citizen Practices." Mexico City.
[37] NatCen Social Research. "British Social Attitudes Survey 32." London, UK.

1.2.2 Relevance of the Thesis

The relevance of this thesis needs to be explicit in order to establish its specific contribution to knowledge. This can be summarised by two aspects. Theoretically, it is relevant because the institutional conditions that shape the way in which the hidden politics of evaluation is manifested have been little explored:

1. Although a large research literature on evaluation studies exists,[38] there has been very little analysis of the relationship between politics and politicians, on the one hand, and policy and evaluation mechanisms on the other.[39] The existing knowledge base has focused on the development of strategies for coping with the political context in which evaluation takes place.[40]

2. Despite the fact that there is large literature on the notions of 'hollowing out' and 'filling in' in relation to governance, none of this has focused specifically on the issue of evaluation as a potential risk-reduction mechanism that may contribute to the 'filling in' of the 'hollowing out' of the state.[41]

3. The effectiveness of policymaking relies on a series of external factors other than just the technical capacity of public agencies. The differences in the resources at disposal among the actors involved, i.e. politicians, operators, can influence the arrangements produced for the design, implementation, and evaluation of policies. In this context, this thesis will focus on generating new knowledge about the role of politics in the use of evaluation as a mechanism for strengthening policymaking and accountability, i.e. to what extent evaluation findings provide more elements for decision-making.

4. The case study comparison will provide evidence about how institutional frameworks influence the results obtained from the

---

[38] See Contandriopoulos, D., and A. Brousselle. 2012. "Evaluation models and evaluation use." *Evaluation* 18(1):61-77, Cunill-Grau, N., and S. M. Ospina. 2012. "Performance measurement and evaluation systems: Institutionalizing accountability for governmental results in Latin America." *New Directions for Evaluation* 2012(134):77-91, Weiss, C. H. 1998a. *Evaluation: Methods for Studying Programs and Policies.* UK.

[39] For example, The SAGE Handbook of Governance that contains the most relevant contemporary discussions about governance does not include policy evaluation as a topic of discussion. Bevir, M. (Ed.). (2011). *The SAGE Handbook of Governance.* London, UK: SAGE.

[40] See Mohan, R., and K. Sullivan. 2006. "Managing the Politics of Evaluation to Achieve Impact." *New Directions for Evaluation* 2006(112):7-23. Bovens, M., P. 't Hart, and S. Kuipers. 2006. "The Politics of Policy Evaluation." in *The Oxford Handbook of Public Policy*, edited by M. Moran, M. Rein, and R. E. Goodin. Oxford, UK: Oxford University Press, Datta, L. E. 2011. "Politics and Evaluation: More Than Methodology." *American Journal of Evaluation* 32(2):273-94, Guenther, J., E. Williams, and A. Arnott. 2010. "The politics of evaluation: evidence-based policy or policy-based evidence?" in *NARU Public Seminar Series.* Darwin, Australia.

[41] See Holliday, I. 2000. "Is the British State Hollowing Out?" *The Political Quarterly* 71(2):167-76, Rhodes, R. A. W. 1994. "The Hollowing Out of the State: The Changing Nature of the Public Service in Britain." Ibid.65:138-51, ibid., Taylor, D. 2005. "Governing Through Evidence: Participation and Power in Policy Evaluation." *Journal of Social Policy* 34(04):601-18.

adoption of evaluation systems.

5. Finally, the focus on evaluation studies offers a lens through which an examination may be made of evaluation studies in Mexico and the UK that may help underline the value of governance as a conceptual framework.

In terms of empirical evidence, this thesis focuses on the implications of the hidden politics of evaluation in the relationships between stakeholders in a particular context. This can provide information about elements that can foster its utilisation, not only for the improvement of policy, but also for the development of arguments vis-à-vis the political debate. Some other identifiable empirical contributions are:

1. Evidence about the influence of stakeholders in the development of evaluation tools, which are not value-free and can predetermine the results.

2. Evidence about the role of sponsors and commissioners of evaluation and the consequences for independence and objectivity.

3. Information about the market of evaluators, the incentives created by these markets and their effect on the quality and reliability of evaluation studies.

4. Evidence about the political use of evaluation, e.g. the manipulation of evaluation results, the use of evaluation results for attacking or discrediting political adversaries, the non-utilisation of evaluation results in the policymaking process due to political reasons, and its connection to the institutional framework.

5. Identification of variables associated with the institutional framework that influences the politics of evaluation can serve as a basis for the discussion of its role in the political debate about the effectiveness of public policy.

The comparison and contrasting of two different evaluation processes can bring new light on the nexus between evaluation and politics and explain *why* this occurs. On the one hand, the UK case is paradigmatic in the sense that NPM reforms emphasised the values of efficiency and efficacy in the public sector although the adoption of evaluation tools has been inconsistent across time. On the other hand, the main outputs of the Mexican evaluation policy are an interesting element of analysis as the Federal Government has devoted important efforts to the institutionalisation of this practice, but little attention has been given to its *actual contribution*. This raises questions about whether

evaluation in this country has permeated the rationale of public organisations or if it has been adopted as another centrally imposed management tool. The specific characteristics of these cases are an additional motivation for the performance of this research, which is described in more detail in the next section.

1.3 How? Theoretical Framework and Methodology

Although both theoretical and methodological frameworks are explained in detail (see Chapters 2 and 3) it is important to briefly introduce their general content. The identification and analysis of the body of knowledge about the hidden politics of evaluation and the process undertaken for the obtaining of empirical evidence are crucial for contextualising the findings obtained.

1.3.1 Generalities of the Theoretical Framework

The literature comprises the existing knowledge about a specific topic that enables the development of new theories. In order for this thesis to represent a theoretically informed and a policy-relevant document, it is important to make explicit the linkages between theory and research.

The theoretical framework of this research is based on three pillars: (1) governance, (2) public policy, and (3) evaluation studies, which when combined provide a structured approach to the hidden politics of evaluation. Governance as 'the new method by which society is governed'[42] is useful for grasping the different transformations that the state has suffered lately, not only in its nature, but also in the appearance of new actors participating in the public sphere.[43] The connection with the use of evaluation tools regards the need of governments for mechanisms of control and legitimisation of public action. This gave evaluation a new meaning because, as Taylor argues, 'From being a contingent instrument of administration, evaluation has become a central legitimating device for a new form of politics.'[44] Under these conditions,

---

[42] Rhodes, R. A. W. 2007. "Understanding Governance: Ten Years On." *Organization Studies* 28(8):1243-64. P. 1246.
[43] Moran, M. 2003. *The British Regulatory State: High Modernism and Hyper-innovation*. Oxford, UK: Oxford University Press.
[44] Taylor, D. 2005. "Governing Through Evidence: Participation and Power in Policy Evaluation."

evaluation has become a mechanism for governments to regain control through the establishment of indicators, evaluation criteria, and methods for their policies. The knowledge developed within the evaluation studies field generates the required tools for a clearer comprehension of its potential uses in the public sector. This thesis aims to provide a broader panorama by focusing on governance as a new paradigm that acknowledges the participation of non-governmental actors and new configurations of power.[45] In this context, evaluation can be perceived as a mechanism to respond to the government crises that have led to the transformation of the state.

The main link between these bigger socio-political shifts and the core of this research is that the changes that governments face have promoted the development and adoption of mechanisms to overcome these new challenges. Policy evaluation sits as a tool directed to improve different areas of public policy. These linkages between theory and research can be summarised as follows:

1. Governance implies the presence of 'inter-organizational networks by interdependence, resource exchange, rules of the game and significant autonomy from the state.'[46] Governments do not take unilateral decisions; hence, there is a strong need to legitimate public action. Evaluation, as an instrument for generating evidence about the effectiveness of the public action, serves this legitimisation purpose, not only through discursive methods, but also by providing theoretically 'objective' evidence about the government's accomplishments.

2. The politicisation of evaluation in the context of the policy process needs to be analysed from an empirical perspective by looking at its effects on the planning and execution of evaluation, as well as on the use of its findings.

3. The use given to evaluation findings is subject to the political influence of the stakeholders of evaluation. Thus, it seems relevant to look at the perennial debate between 'good policy' and 'good politics' in the context of the two selected countries.

4. Finally, the hidden politics of policy evaluation as the core of this research is a factor of the policy process that the theory has only partially explored. This research aims to focus on those areas in which there is political influence, as well as to identify the positive outcomes of acknowledging the political nature of evaluation.

Journal of Social Policy 34(04):601-18. P. 602.

[45] Cabrero, E., and Nava, G. (Eds.). (1999). *Gerencia Pública Municipal: Conceptos Básicos y Estudios de Caso*. México, D.F.: Miguel Ángel Porrúa-CIDE.

[46] Rhodes, R. A. W. 1997. *Understanding Governance: Policy Networks, Governance, Reflexivity and Accountability*. Buckingham, UK: Open University Press. P. 15.

1.3.2 Overview of the Methodology

The elusiveness of the political factor existing in evaluation highlights the need for methodological tools that can capture the influence of contextual variables and the interaction of different stakeholders in the process.[47] Recently, there has been an important effort from scholars to explain how politics can threaten evaluation in terms of its integrity. It has been acknowledged that conditions of independence and impartiality need to be preserved in order for evaluation to be considered credible and robust.[48] However, the political component of evaluation implies a broader range of considerations that concern the level of influence that different actors have in the process and outcomes of evaluation.[49] The development of a comparative study based on a qualitative approach will allow making inferences about the political nature of evaluation in two contexts in which evaluation is ruled by different procedures and institutions, i.e. the existence of a central agency devoted to the evaluation *vis-à-vis* a model when government departments adopt this function. The focus of the research is the comparison of four case studies (two per country) that belong to the health sector. The main purpose is to study a set of variables—purpose, resources, process, outputs, and outcomes—to explore the influence of the institutional framework in the politics of evaluation. The main tool for obtaining evidence, aside from documentary analysis, is the performance of circa 40 semi-structured interviews to stakeholders of evaluation.

An important point to discuss in this introduction is the rationale for the selection of the case studies. Different institutional arrangements have shaped the way in which evaluation systems are developed and implemented in both countries, which are the result of their particular political and administrative conditions.[50] Mexico, as an emergent democracy, offers an interesting

---

[47] See Schwartz, R. (1998). The Politics of Evaluation Reconsidered. *Evaluation, 4*(3), 294-309.

[48] Lipsky, D. B., R. L. Seeber, A. C. Avgar, and R. M. Scanza. 2007. "Managing the Politics of Evaluation: Lessons from the Evaluation of ADR Programs." Pp. 116-29 in *Proceedings of the Fift-Ninth Annual Meeting of the Labor and Employment Relations Association*.

[49] This can be observed in the level of influence that operators can have in the selection of methods. See: Chelimsky, E. 2012. "Valuing, Evaluation Methods, and the Politicization of the Evaluation Process." *New Directions for Evaluation* 2012(133):77-83.

[50] *Appendix A* provides an overview about the main differences between the political systems of these countries.

perspective of how the use of evaluation tools has been crucial for the establishment of political balance. The promotion of evaluation in this country has been linked to the need for institutionalising these efforts within the Federal Government.[51] The case of the UK provides a different lens, through which the role of evaluation in policymaking and for accountability in the context of a more developed and stable democracy may be understood.[52] Although the socio-political differences between these countries are not the focus of this thesis, it is important to understand how rules and institutions operate for the performance of evaluations. An interesting issue, for instance, is looking at the relationship between the audit institutions in these countries and the legislative body.[53]

The selection of cases within the health sector responds to the fact that there is an important range of evaluation outputs susceptible to analysis. In both countries, health and education are among the most developed policy areas in terms of evaluation.[54] Regardless of how policy is implemented in each country, these cases allow the observation of how health policy is evaluated and how 'success' is defined. Also, the values that guide evaluation, the constraints imposed by the institutional framework, and the use given to findings are interesting variables to contrast between countries.

In terms of the methodological process, this research can be divided into three main phases. The first one was the construction of the theoretical framework that required bibliographical research to identify the most relevant concepts that sustain the hypothesis. This stage allowed operationalising the concept of

---

[51] Castro, M. F., G. Lopez-Acevedo, G. Beker Busjeet, and X. Fernandez-Ordonez. 2009. "Mexico's M&E System: Scaling Up from the Sectoral to the National Level." Washington, D.C., US: The World Bank.

[52] Heady, F. 2001. *Public Administration: A Comparative Perspective*. New York, US: Marcel Dekker.

[53] The case of the UK was studied by the OECD who describes the relationship between the NAO and the PAC. See: OECD. 2002. "Relations Between Supreme Audit Institutions and Parliamentary Committees." in *Sigma Papers*.

[54] In the case of the UK, for example, the National Institute for Health and Care Excellence (NICE) and the Education Endowment Foundation (EEF) are organisations directed to generate evidence for policymaking. In the Mexican case, health and education sector are also prominent areas where evaluation is an important asset. The National Institute of Public Health (INSP, its acronym in Spanish) and the National Institute for the Evaluation of Education (INEE, its acronym in Spanish) are constantly producing evidence to inform public decision.

evaluation in key variables for the analysis. In addition, it involved the characterisation of the institutional framework for evaluation operating in both countries. The second phase comprised the collection and systematisation of data, through documentary analysis and interviews. The last phase was the analysis of the data that allowed observing the influence of politics in the evaluation process. Here, case studies were compared through the lenses of the thematic framework proposed in this thesis.

This overview of the theoretical and methodological framework shows that the core of the research concerns the institutional framework and the political nature of evaluation. The nuances of this discussion will be illustrated through the findings obtained from the case studies and their relation to the theory. The next chapter is dedicated to the development of the theoretical framework by looking at key elements of the literature about governance, public policy, and evaluation studies.

Chapter 2. Theoretical Framework

This chapter presents the most relevant theoretical discussions about the hidden politics of evaluation that sustain the core argument of this thesis. The connection between evaluation and politics demands a broader development than the one found in the evaluation studies field.[55] It requires widening the perspective to elucidate how evaluation has gained relevance in the reform of the public sector and in the transformations of the state. This literature review addresses three major arguments:

1. The transition from government to governance promoted the adoption of monitoring, evaluation, and audit systems to strengthen policymaking and to support accountability by giving stakeholders more elements to participate in the political arena, e.g. by promoting changes, preserving the *status quo,* etc.
2. The second argument concerns the role of these activities as an aid for stakeholders to interact in a context of uncertainty and limited resources.
3. Both evaluation and auditing are underpinned by a political logic that neither can nor should be ignored, as it affects both the process and the outputs of these activities.

This theoretical framework relies on three pillars: governance, public policy, and evaluation studies. The first section is dedicated to the key discussions of governance, public policy, and underlining the linkage between evaluation and performance management. The second section presents the theoretical grounds of evaluation and its institutionalisation. The final part discusses the existing knowledge about the political nature of evaluation. This chapter aims to highlight the way in which the political context fits into the transforming processes of the modern state and how it has promoted evaluation as a mechanism for strengthening decision-making, reducing uncertainty, controlling risk, and legitimising public action.

2.1 Governance, Public Policy, and Performance Management

The purpose of this section is to tease out those elements of governance, public policy, and performance management that relate to the hidden politics of evaluation. The central argument of this thesis is that the institutional

---

[55] Palumbo, D. J. 1987. *The Politics of Program Evaluation.* California, US: SAGE, Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106.

framework affects the way in which politics and evaluation interact. Uncovering the hidden politics of evaluation demands taking a step back to identify the role of evaluation in governance and, more specifically, within policymaking. Evaluation participates in the transformation of the state by providing evidence for different stakeholders to act according to their interests, while contributing to democratic values such as participation and transparency.[56] This transformation entails the establishment of new cooperative networks for the identification of goals and causal effects of public policy, as it occurs, for example, in the different ways in which social programmes benefit people.[57]

Beyond its merely instrumental use, there are important consequences to acknowledge about evaluation systems. In a context in which governments have lost the monopoly of public action, evaluation systems have inevitably affected the relationships between participants, specifically in the control over the evaluation processes, i.e. between commissioners and evaluators.[58] The fragmentation of power in the public sphere has fostered the use of evaluation to reduce uncertainty and to legitimate public action.[59] Considering the characteristics of the modern state (reviewed in this section), evaluation plays an important role in the development of new configurations of power for governments to respond to the demands of stakeholders.[60]

2.1.1 Governance

The term 'governance'[61] explains the transformation of the state due to the challenges imposed by the context and its impact 'on present and future levels of human wellbeing, political stability, and democratic vitality.'[62] As the need for

---

[56] Stame, N. 2006. "Governance, Democracy and Evaluation." *Evaluation* 12(1):7-16.

[57] See: Hasenfeld, Y., K. Hill, and D. Weaver. "A Participatory Model for Evaluating Social Programs." San Francisco, CA: The James Irvine Foundation.

[58] Themessl-Huber, M. T., and M. A. Grutsch. 2003. "The Shifting Locus of Control in Participatory Evaluations." *Evaluation* 9(1):92-111.

[59] Pollitt, C. 2006. "Performance Information for Democracy: *The Missing Link*?" Ibid.12:38-55.

[60] House, E. R. Ibid."Democracy and Evaluation."119-27.

[61] See also: Pierre, J., and B. G. Peters. 2000. *Governance, Politics and the State*. Basingstoke, UK: Palgrave Macmillan, Rhodes, R. A. W. 1997. *Understanding Governance: Policy Networks, Governance, Reflexivity and Accountability*. Buckingham, UK: Open University Press, Richards, D., and M. J. Smith. 2002. *Governance and Public Policy in the United Kingdom*. Oxford, UK: Oxford University Press.

[62] Duit, A., and V. Galaz. 2008. "Governance and Complexity—Emerging Issues for Governance Theory." *Governance* 21(3):311-35. P. 311.

public intervention is dynamic, governments have been forced to adapt themselves to new conditions, i.e. economic crises and more active civil society, and to rethink how policy is designed, implemented, and evaluated.[63]

The transition from government to governance relates to the discussions about the role of governments after World War II. For the decades that followed, public sectors were 'increasingly seen as rigid and bureaucratic, expensive and inefficient.'[64] As a response, a series of reforms of the public sector took place in countries like the UK, New Zealand, and Australia,[65] the core of which was …

> … to restrain expenditure […] increasing the productivity of public services –'getting more for less', or raising efficiency […] moulding public services more closely in accordance with the convenience and preferences of their users […] strengthening 'transparency', usually on the assumption that transparency is a feature of democratic governance and that more of it should strengthen public confidence in the workings of the State.[66]

Most of the drivers of these reforms were related to crises of efficiency and legitimacy.[67] In Mexico, for example, citizens' perception about corruption relates to the legitimisation crisis faced by the state.[68] An important component of the Federal Government reform plans has been the adoption of instruments that could improve the effectiveness of public policy and accountability.[69] These crises highlighted the need for introducing more rationality into policymaking and for new mechanisms of accountability.[70] This provoked a shift in the conceptualisation of the state,[71] in terms of how control mechanisms were

---

[63] Howlett, M. 2000. "Managing the "hollow state": procedural policy instruments and modern governance." *Canadian Public Administration-Administration Publique Du Canada* 43(4):412-31.
[64] Pierre, J., and B. G. Peters. 2000. *Governance, Politics and the State*. Basingstoke, UK: Palgrave Macmillan. P. 5.
[65] Bevir, M., R. A. W. Rhodes, and P. Weller. 2003. "Traditions of Governance: Interpreting the Changing Role of the Public Sector." *Public Administration* 81(1):1-17.
[66] Pollitt, C. 1999. "The Changing Context." in *Performance or compliance?: Performance Audit and Public Management in Five Countries*, edited by C. Pollitt, X. Girre, J. Londsdale, R. Mul, H. Summa, and M. Waerness. Oxford, UK: Oxford University Press. P. 53-54.
[67] Cabrero, E., and G. Nava (Eds.). 1999. *Gerencia Pública Municipal: Conceptos Básicos y Estudios de Caso*. Mexico City: Miguel Ángel Porrúa-CIDE.
[68] Bailey, J., and P. Paras. 2006. "Perceptions and Attitudes about Corruption and Democracy in Mexico." *Mexican Studies* 22(1):57-82.
[69] Former President Ernesto Zedillo (1995-2000) established the Public Administration Modernisation Programme which recognised that Mexican citizens demanded 'the continuous improvement of government services, more transparency and honesty in the government's actions and in the use of public funds, also a more comprehensive accountability process…' "Public Administration Modernisation Programme 1995-2000." Mexico City, Mexico.
[70] The World Bank. 1992. *Governance and development*. Washington, D.C., US: The World Bank.
[71] Fukuyama, F. 2013. "What Is Governance?" *Governance* 26(3):347-68.

designed and implemented.[72] Efficiency became a core value of governance and it fostered measures such as administrative simplification, results-oriented government, and a focus on outcomes.[73] This intended 'to improve management performance by freeing them from centralized controls over the management of resources.'[74] It also demanded a new configuration of processes within government, the establishment of more direct relationships with non-governmental actors, and the implementation of new models of scrutiny that can adapt to this new scenario.[75] According to Rhodes, governance …

> … refers to a change in the meaning of government, referring to a new process of governing […] It refers to self-organizing, inter-organizational networks characterized by interdependence, resource exchange, rules of the game and significant autonomy from the state.[76]

The core element of this definition is the identification of networks composed by new actors that 'ignore conventional public/private sector boundaries [and] link in relations of mutual dependence.'[77] This demanded new schemes of operation, based on reciprocity and cooperation which gave evaluation a new meaning within governance models.[78] Under these conditions, private and public actors interact in the policy process.[79] However, their level of engagement varies according to the context. For example, in the UK the British Medical Association is constantly approaching the government to make it aware of its concerns.[80] In contrast, until very recently, the Mexican Federal Government has faced more contesting voices in relation to its policies.[81]

---

[72] Considine, M. 2002. "The End of the Line? Accountable Governance in the Age of Networks, Partnerships, and Joined-Up Services." *Governance: An International Journal of Policy, Administration, and Institutions* 15(1):21-40.

[73] Osborne, D., and P. Plastrik. 1998. *Banishing Bureaucracy: The Five Strategies for Reinventing Government.* New York, US: Plume.

[74] Aucoin, P. 2012. "New Political Governance in Westminster Systems: Impartial Public Administration and Management Performance at Risk." *Governance* 25(2):177-99. P. 178.

[75] Johnson, C., and C. Talbot. 2007. "The UK Parliament and performance: challenging or challenged?" *International Review of Administrative Sciences* 73(1):113-31.

[76] Rhodes, R. A. W. 1997. *Understanding Governance: Policy Networks, Governance, Reflexivity and Accountability.* Buckingham, UK: Open University Press. P. 16.

[77] Moran, M. 2003. *The British Regulatory State: High Modernism and Hyper-innovation.* Oxford, UK: Oxford University Press. P. 12.

[78] Hanberger, A. 2006. "Evaluation of and for Democracy." *Evaluation* 12(1):17-37.

[79] Héritier, A., and D. Lehmkuhl. 2008. "The Shadow of Hierarchy and New Modes of Governance." *Journal of Public Policy* 28(1):1-17.

[80] See for example: ITV News. 2015. "British Medical Association chief warns next government against introducing charges on the NHS." London, UK.

[81] A clear example of this was the strong criticism that former president Calderón (2006-2012) received in relation to the 'drug war' he started and that led to a substantial increasing of the

The characteristics of the governance model can also be observed in the inclusion of concepts like transparency and accountability, now constantly merged in the public discourse.[82] Evaluation is seen as a mechanism for giving voice to those actors usually marginalised and for opening new spaces for debating the effectiveness of public policy.[83] The basis on which stakeholders can interact, negotiate, and participate have been modified because 'the boundaries between public, private and voluntary sectors became shifting and opaque.'[84]

Evaluation processes gain importance because of this shift in the role of the state. On the one hand, the scarcity of resources and the multiplicity of social demands require policymakers and politicians to make decisions based on robust and credible evidence.[85] On the other hand, the presence of new actors in the public dynamics implies establishing new instruments for evaluating the results delivered by governments and to hold them accountable.[86]

The work of Flinders is useful for making the connection between evaluation and governance clearer. This author identifies four core variables that can guide this discussion: *control, coordination, accountability*, and *power*.[87] This framework helps grasp the concept of governance, transcending its adoption as a pre-established model and seeing it as an evolution of the conceptualisation of the public sector.[88] Also, it allows highlighting the relevance of the institutional framework by looking at the rules for the distribution of responsibilities for the conduction of the policy process.

---

violence in the country. See: Feeney, M. 2012. "Leaving Office Calderón Leaves Bloody Drug War Legacy, Tens of Thousands Dead and Missing." in *Hit and Run*.

[82] These concepts for example can be found in the Conservative Party Manifesto in the UK, as well as in the National Development Programme of the Mexican government. Presidencia de la República. "Plan Nacional de Desarrollo 2013-2018." Mexico City, The Conservative Party. "The Conservative Manifesto 2015."

[83] Cook, J. R. 2014. "Using Evaluation to Effect Social Change: Looking Through a Community Psychology Lens." *American Journal of Evaluation* 36(1):107-17.

[84] Rhodes, R. A. W. 2007. "Understanding Governance: Ten Years On." *Organization Studies* 28(8):1243-64. P. 1246.

[85] Rotberg, R. I. 2014. "Good Governance Means Performance and Results." *Governance* 27(3):511-18.

[86] Aucoin, P. 2012. "New Political Governance in Westminster Systems: Impartial Public Administration and Management Performance at Risk." Ibid.25(2):177-99.

[87] Flinders, M. 2002. "Governance in Whitehall." *Public Administration* 80(1):51-75.

[88] Andrews, M. 2010. "Good Government Means Different Things in Different Countries." *Governance* 23(1):7-35.

While in the past government was considered a key actor, in governance it becomes a player in the political arena, retracted from different spheres of the public sector. The Organisation for Economic Cooperation and Development (OECD) recognises that 'governments have progressively withdrawn from commercial activities, ownership of industries and service provision.'[89] This has created new spaces of power that have become at dispute with different stakeholders. The adoption of more indirect control mechanisms adopted by governments resonates with the discussions about depoliticisation. According to Flinders and Buller, depoliticisation can be understood as:

> The range of tools, mechanisms and institutions through which politicians can attempt to move to an indirect governing relationship and/or seek to persuade the demos that they can no longer be reasonably held responsible for a certain issue, policy field or specific decision.[90]

This does not imply an absence of politics. On the contrary, it refers to the adoption of new schemes of organisation that exceed the governmental sphere, as it was traditionally conceptualised—rigid and inflexible structures, vertical hierarchies, and direct control lines.[91] Hay coincides with this position:

> Depoliticisation is not about less politics, but about a displaced and submerged politics – a politics occurring elsewhere, typically beyond sites and arenas in which it is visible to nonparticipants and hence amenable to public – perhaps even democratic – scrutiny.[92]

Here, this concept facilitates comprehending how governments exercise control. Depoliticisation 'reduces government responsibility for policy while also leaving it less subject to political discretions, deliberations and interventions.'[93] However, as Fawcett and Marsh recognise, there is little notion about how this discretion is used and how interests are transformed into inputs and outputs of the political process.[94] This is pertinent for discussing evaluation in the context

---

[89] OECD. 2005. *Modernising Government: The Way Forward*. Paris, France: OECD. P. 16.
[90] Flinders, M., and J. Buller. 2006. "Depoliticisation: Principles, Tactics and Tools." *British Politics* 1(3):293-318. P. 295-296.
[91] See Wood, M., and M. Flinders. 2014. "Rethinking depoliticisation: beyond the governmental." *Policy & Politics* 42(2):151-70.
[92] Hay, C. Ibid."Depoliticisation as process, governance as practice: What did the 'first wave' get wrong and do we need a 'second wave' to put it right?":293-311. P. 302.
[93] Kuzemko, C. Ibid."Politicising UK energy: what 'speaking energy security' can do."259-74. P. 261.
[94] Fawcett, P., and D. Marsh. Ibid."Depoliticisation, governance and political participation."171-88.

of what Wood and Flinders call *government depoliticisation.*[95] Evaluation is associated with the application of technical instruments to analyse different dimensions of public policy. This technocratic label, in a sense, establishes some distance from government and it delegates this function to other stakeholders (think tanks, academicians, etc.) who can provide an 'unbiased' judgement of policy and can—or at least attempt to—reduce the risk of politicisation.[96]

The adoption of new mechanisms of governing produces new arrangements among stakeholders, in which evaluation and audit can play an important role. The obsolescence of traditional control mechanisms is perhaps one of the reasons that promoted the rise of these systems.[97] The relationship between the allocation of resources and the use of incentives, for example, is a scheme clearly focused on the measurement of performance.[98] This affects the levels of control because of the fragmentation of the bureaucratic body, the emergence of autonomous agencies and the appearance of non-governmental organisations (NGOs).[99] This decentralisation shows that 'the government swapped direct for indirect controls and central departments are no longer either necessarily or invariably the fulcrum of a network.'[100]

This can be perceived, for example, in the creation of specialised agencies and decentralised bodies.[101] According to Talbot, this is the result of public-reform trends that impose new challenges in terms of control.[102] There is an explicit

---

[95] For them, this 'focuses on the transfer of issues from the governmental sphere to the public sphere through the 'delegation' of those issues by politicians to arm's-length bodies, judicial structures or technocratic rule-based systems that limit discretion.' Wood, M., and M. Flinders. Ibid."Rethinking depoliticisation: beyond the governmental."151-70. P. 165.

[96] This resonates with Moran's ideas about the rise of regulatory agencies 'free from political partisan in an effort to guide policy by technocratic imperatives rather than by the outcomes of partisan, majoritarian politics.' Moran, M. 2003. *The British Regulatory State: High Modernism and Hyper-innovation.* Oxford, UK: Oxford University Press. P. 12.

[97] See: Kudo, H. 2003. "Between the 'Governance' Model and the Policy Evaluation Act: New Public Management in Japan." *International Review of Administrative Sciences* 69(4):483-504.

[98] Hood, C. 1991. "A public management for all seasons?" *Public Administration* 69:3-19.

[99] The case of the Care Quality Commission discussed in this thesis (see Chapter 4) illustrates this situation in terms of the complexity derived from new models of control in organisations like arm's–length bodies.

[100] Rhodes, R. A. W. 2007. "Understanding Governance: Ten Years On." *Organization Studies* 28(8):1243-64. P. 1245.

[101] For further reference see: Greve, C., M. Flinders, and S. Van Thiel. 1999. "Quangos—What's in a Name? Defining Quangos from a Comparative Perspective." *Governance* 12(2):129-46.

[102] Pollitt, C., and C. Talbot. 2004. *Unbundled Government: A critical analysis of the global trend to agencies, quangos and contractualisation.* London, UK: Routledge.

attempt of decentralising and maintaining some distance between central government and key policy areas, despite the linkage being still observable through the mechanisms of control adopted, e.g. by giving technical but not financial autonomy. Therefore, the boundaries between public and private spheres become even more blurred:

> In the last thirty years, the capacity of the state to control or direct society and the extent to which institutions of central government retain a monopoly on political power have become a much more contested issue [...] governance has become the established concept within political science which is used to portray the changing nature of the state in recent times.[103]

Governance 'enables [...] to understand better the role that non-governmental actors play in producing policy outcomes.'[104] Evaluation and auditing represent control mechanisms through the use of evidence and the establishment of criteria for measuring performance, a central element of the NPM approach.[105] As Boardman suggests, governance highlights the relevance of evaluating 'both processes and outcomes' to have more instruments for decision-making.[106] The question that arises here is: *How should control be allocated in terms of the design and implementation of evaluation and audit systems*? This means: *Who will evaluate? Who will be evaluated? Whose criteria will predominate?* These are all political questions that can be addressed through the establishment of rules and agreements, but they require a profound understanding of how these systems interact with the political context.

Coordination can be understood as the capacity of aligning an actor's behaviour towards a desired outcome. In governance, the accomplishment of goals depends on the capacity of coordination. Actors are required to interact with acknowledging that 'The image of hierarchical relationship between government and citizenry is displaced by the idea of multiple parallel spaces in

---

[103] Richards, D., and M. J. Smith. 2002. *Governance and Public Policy in the United Kingdom*. Oxford, UK: Oxford University Press. P. 13.

[104] Pierre, J., and B. G. Peters. 2000. *Governance, Politics and the State.* Basingstoke, UK: Palgrave Macmillan. P. 30.

[105] Hansson, F. 2006. "Organizational Use of Evaluations: Governance and Control in Research Evaluation." *Evaluation* 12(2):159-78.

[106] Boardman, C. 2014. "Assessing Governance: The Importance of Evaluating Policy Outcomes in National Mission Areas." *Governance* 27(3):519-26. P. 523.

which power is encountered and negotiated.'[107] Actors (with uneven capabilities and power) need evidence to sustain their positions and agendas and the alignment of their behaviour is difficult to accomplish.[108] The means through which political goods can be measured demand incorporating the logic of multiple stakeholders who require risk-reduction tools to promote these agendas, based on robust information.[109]

Networks between actors have promoted the discussion about the rise of new mechanisms of coordination.[110] This leads to reflecting upon how this 'creates tensions between institutions' new and traditional roles.'[111] For example, Painter observed that the 2010 Coalition's reform project was:

> An adaptation to emergent realities of 'networked community governance', as the centre of gravity moves from state not only to markets but towards civil society, a context in which services are increasingly 'co-produced' through a variety of informal and formal configurations, rather than being provided hierarchically through state bureaucracies.[112]

broader spectrum of stakeholders shows how evaluation and auditing relate to the identification of results produced by these networks.[113] Also, it leads to reflecting on their role in the control and implementation of evaluation systems. It imposes the challenge of determining the use that will be given to evaluation, ranging from a tool for evidence-generation to a central element in the construction of the political discourse.[114] Cooperation, therefore, regards also

[107] Kennett, P. (Ed.). 2008. *Governance, Globalization and Public Policy*. Cheltenham, UK: Edward Elgar Publishing Ltd. P.4.

[108] This complexity can be observed in the discussions about joined-up government and its role in the context of the public sector reform. Ling, for example, analysed the case of the UK and observed that the adoption this model entails reflecting about 'the way in which intra-state relationships are managed opens up the state to different forms of participation by different interests.' This evinces the fact that the rise in the number of actors involved requires looking at the mechanisms that will norm their interaction. Ling, T. 2002. "Delivering Joined-Up Government in the UK: Dimensions, Issues and Problems." *Public Administration* 80(4):615-42. P. 639.

[109] Rotberg, R. I. 2004. "Strengthening governance: Ranking countries would help." *The Washington Quarterly* 28(1):71-81.

[110] Risse, T. 2012. "Governance in Areas of Limited Statehood." in *The Oxford Handbook of Governance*, edited by D. Levi-Faur. Oxford, UK: Oxford University Press.

[111] Pierre, J. Ibid."Governance and Institutional Flexibility." P. 187.

[112] Painter, C. 2012. "The UK Coalition government: Constructing public service reform narratives." *Public Policy and Administration* 20(1):3-20. P. 4.

[113] Hertting, N., and E. Vedung. 2012. "Purposes and criteria in network governance evaluation: How far does standard evaluation vocabulary takes us?" *Evaluation* 18(1):27-46.

[114] Segerholm, C. 2003. "Researching Evaluation in National (State) Politics and Administration: A Critical Approach." *American Journal of Evaluation* 24(3):353-72.

the creation of incentives that shape the behaviour of stakeholders. Under this idea, evaluation policies require cooperation not only for the implementation of evaluation instruments, but also for the use/dissemination of findings and for the overall credibility of the system.

The third variable discussed by Flinders is *accountability.* This concept has provoked the reconfiguration of processes in the logic of politicians and practitioners, as well as the use of incentives to shape individuals' behaviour.[115] Considering the decomposition of the term proposed by Schedler, there are two elements of accountability. First, *Answerability,* seen as the obligation of public servants and politicians to inform systematically about their decisions and actions. Both in the Executive and Legislative branches, accountability systems produce information about the progress made in policymaking and about the effective use of public funds. Nonetheless, as Romzek and Dubnick reflected on it, the idea of *answerability* is more complex because it entails coping with the expectations created around public organisations.[116]

Enforcement, the second element, is the capacity of sanctioning any public servant who infringes the law or who does not fulfil his/her obligations.[117] Evaluation, as an instrument for explaining and justifying public action, can be susceptible to politics because of this component. Accountability has given evaluation a window of opportunity to become more imbued in the logics of public organisations.[118] At the same time, it imposes the challenge to reflect upon the use of evaluation data in accountability systems.[119]

Evaluation and audit systems in the context of governance allow 'framing our expectations of government programs, policies, and public services.'[120] This

---

[115] For further knowledge about incentives within accountability see: Ackerman, J. M. 2005. "Social Accountability in the Public Sector. A Conceptual Discussion." edited by Social Development Department. Washington, D.C., US: The World Bank.

[116] Romzek, B. S., and M. J. Dubnick. 1987. "Accountability in the Public Sector: Lessons from the Challenger Tragedy." *Public Administration Review* 47(3):227-38.

[117] Schedler, A. 2008. *¿Qué es la rendición de cuentas?* Mexico City.

[118] Carman, J. G. 2009. "Nonprofits, Funders, and Evaluation: Accountability in Action." *The American Review of Public Administration* 39(4):374-90.

[119] Benjamin, L. M. 2008. "Evaluator's Role in Accountability Relationships: Measurement Technician, Capacity Builder or Risk Manager?" *Evaluation* 14(3):323-43.

[120] Chouinard, J. A. 2013. "The Case for Participatory Evaluation in an Era of Accountability." *American Journal of Evaluation* 34(2):237-53. P. 248.

comprises identifying and disseminating the achievements made by public organisations, but also the changes that policy has produced in the population, e.g. increased quality of life, which can be highly politically profitable. Therefore, evaluation can be used for internal purposes, i.e. organisational learning, as well as externally, i.e. political recognition.[121]

The contribution of evaluation to democracy is linked to its capacity to support accountability by providing evidence for informing and justifying actions, and also for enlightening the policy process.[122] These instruments represent an important source of political legitimacy by making explicit commitment to democratic values like transparency and accountability, particularly in contexts with deficits of credibility and trust in the state.[123]

Although governance represents a more fertile ground for the establishment of an accountability/evaluation relationship, the institutional framework needs to incorporate safeguards that can give legitimacy to this bond by securing the conditions for the performance of evaluations, the validity of the methods used, and the appropriate dissemination of the findings.[124] However, the effectiveness of this binomial is not automatic; conflicts may arise because of the different nature of these concepts (evaluation and accountability).[125]

This finds echoes in Dubnick's work about 'the promises of accountability'.[126] According to this author, 'We cannot – and should not – continue to rely on the

---

[121] This finds echo in the ideas of Saunders about the *use* and *usability* of evaluation, this is based 'on a distinction between the organizational and political context of use on the one hand, with a particular concern with the capacity of stakeholders within an organization to make use of the knowledge resources made available through evaluation along with the systemic characteristics of the environment […] usability refers to the dimensions of evaluation design, within the power of evaluators to affect, which are likely to inhibit or enhance the chances of evaluation output being used.' Saunders, M. 2012. "The use and usability of evaluation outputs: A social practice approach." *Evaluation* 18(4):421-36. P. 433.
[122] Picciotto, R. 2015. "Democratic evaluation for the 21st century." Ibid.21(2):150-66.
[123] Hanberger, A. 2006. "Evaluation of and for Democracy." Ibid.12(1):17-37.
[124] Hoefer, R. 2000. "Accountability in Action?: Program Evaluation in Nonprofit Human Service Agencies." *Nonprofit Management and Leadership* 11(2):167-77.
[125] Rutman, L. 1990. "Serving Management and Accountability: A Conflict for Program Evaluation? ." *Evaluation Practice* 11(1):39-43.
[126] This author states that 'the current problems associated with accountability in the Extended State are rooted in a widely held but unsubstantiated belief in the capacity of accountability mechanisms alone to bring about the three things we require of our government today— efficient control, democratic legitimacy, and effective performance.' Dubnick, M. J., and H. G. Frederickson. 2011b. "Public Accountability: Performance Measurement, the Extended State, and the Search for Trust." Washington, D.C., US: Kettering Foundation. P. 5.

assumed relationship between accountability and performance that underlies much of the NPM reform agenda.'[127] While it is true that evaluation and performance systems can be important inputs for accountability, the construction of indicators and evaluation tools, the arrangements for the conduction of the process, and for the use and dissemination of findings are aspects that need to be understood carefully, particularly in terms of how this information will allow the identification of responsibilities and the establishment of sanctions. The connection between accountability and performance, as Dubnick recognises, is an assumption that needs to be studied in detail.[128] In particular, it is worth reflecting on how evaluation can help accountability to keep its promises— borrowing Dubnick's term.[129] This means: *To what extent can evaluation be a vehicle for generating evidence about a government's performance, by actually producing reliable and useful information?* And more importantly: *How is evaluation data translated into meaningful evidence about an organisation or an individual's performance?* Without reflecting on this, evaluation's contribution to accountability faces the risk of becoming a purposeless activity.

Dubnick's work is also useful for establishing a link between the 'accountability space' and the 'evaluation space'.[130] Following this author's ideas, the development of relationships between those who claim for evaluation and those who are evaluated creates a space of interaction that is susceptible to control. This makes sense, for example by looking at the accountability function of evaluation. The control of the process of evaluation would be linked to the interests of stakeholders in shaping the outputs that will feed into accountability systems. As a result, uncovering the hidden politics of evaluation implies also identifying how this 'evaluation space' is allocated among

---

[127] Dubnick, M. J. 2005. "Accountability and the Promise of Performance: In Search of the Mechanisms." *Public Performance & Management Review* 28(3):376-417. P. 402.

[128] Dubnick, M. J., and H. G. Frederickson. 2011a. *Accountable Governance: Problems and Promises*. London: Routledge.

[129] Dubnick's ideas about 'the promises of accountability' are also linked to his discussion about the concept of accountability and the relevance of acknowledging how different actors might misuse it. Dubnick, M. J. 2002. "Seeking Salvation for Accountability." in *Annual Meeting of the American Political Science Association*.

[130] —. 2011. "More over Daniel: We need some 'accountability space'." *Administration & Society* 43(6):704-16.

stakeholders in terms of *who* will determine the design, conduction, and dissemination of this activity. Moreover, this gains importance in governance when evaluation is no longer a purely governmental affair and other stakeholders have a more active role in this activity.[131] The identification of this 'evaluation space' can be seen, following Dubnick's ideas about accountability, as the output of the interaction between different actors, e.g. evaluators and commissioners, in the collective construction of the meaning of evaluation, of the criteria that will be established, and more importantly, of the way in which its outputs will feed into other processes within and outside the government's sphere. Here, the locus of the politics of evaluation would be on the disputes for accessing this space.

The last variable identified by Flinders is *power.* Incorporating new actors in the public arena has redefined the way in which power is allocated.[132] This dissemination of power is incompatible with former hierarchical structures that have 'given way to complex networks in which power is a fluid positive-sum concept and a resource held by all actors.'[133] Therefore, power depends on the actors' capacity to obtain and exchange resources, to share information, and to promote their interests. The political arena is a continuous struggle of power and a constant reallocation of it.[134] Interests represented in public policy reflect those issues that get into the public agenda and whoever are the winners and losers after a decision is made. This helps understanding the ability of each actor to pursue an agenda and how the institutional framework affects the allocation of power. Harrison et al. reflect upon the multiple interests involved in evaluation by looking at the UK's health sector:

> One might see government pushing for measures of efficiency and cost-effectiveness (what is achieved per unit of resource input) and the medical profession and other provider organizations advocating instead

---

[131] In the UK, for instance, the Institute for Government produces systematically a range of studies about different policy issues.

[132] Bovaird, for instance, reflects about the idea the regulation of power among stakeholders implies reflecting upon the kind of mechanisms that need to be put in place for this purpose, i.e. checks and balances. See: Bovaird, T. 2005. "Public governance: balancing stakeholder power in a network society." *International Review of Administrative Sciences* 71(2):217-28.

[133] Flinders, M. 2002. "Governance in Whitehall." *Public Administration* 80(1):51-75. P. 69.

[134] Mintzberg, H. 1985. "The organization as political arena." *Journal of Management Studies* 22(2):133-54.

measurement of population health needs and of growth in services.[135]

This shows how conflicting interests might clash and that the outcome of this grating will affect the design and course of an evaluation (see Section 2.3.2), but in addition it shows how the allocation of power in governance has shifted. This resonates with what Sbragia calls 'the two faces of the state'.[136] She recognises: 'There has been a shift of power within the state—some have increased their power at the expense of others.'[137] Evaluation and audit systems are entangled in this phenomenon because these activities can show that political decisions taken might have been suboptimal.[138] The recognition of the political nature of evaluation as the identification of the relationships of power established as a consequence of this activity[139] acknowledges that power is constantly in conflict and that evaluation must be 'seen as *socially located* and understood as *politically contested.'*[140]

In this new configuration of power, evaluation offers stakeholders the capacity to establish reciprocal relationships, based on the idea of 'objective' and legitimate evidence. Although there is no such thing as value-free evaluation, stakeholders can agree on the definition of criteria for evaluating policy and interacting on this basis. The values of rationality and neutrality given to evidence are key to understanding the implications of evaluation as a power tool in policymaking.[141] This resonates with Zapico-Goñi's ideas:

---

[135] Harrison, S., D. J. Hunter, and C. Pollitt. 1990. *The dynamics of British health policy*. London,UK: Unwin Hyman. P. 121.

[136] This refers to the distinction between those actors that are essential for generating benefits from those necessary as 'builder of markets'.

[137] Sbragia, A. M. 2000. "Governance, the State, and the Market: What Is Going On?" *Governance* 13(2):243-50. P. 247.

[138] Heider, C. 2011. "A Conceptual Framework for Developing Evaluation Capacities: Building on Good Practice." in *Influencing Change. Building Evaluation Capacity to Strengthen Governance*, edited by R. C. Rist, M. H. Boily, and F. Martin. Washington, D.C., US: The World Bank.

[139] See: Chelimsky, E. 1995. "The Political Environment of Evaluation and What it Means for the Development of the Field: Evaluation for a New Century: A Global Perspective." *American Journal of Evaluation* 16(3):215-25, Melkers, J., and D. Roessner. 1997. "Politics and the Political Setting as an Influence on Evaluation Activities: National Research and Technology Policy Programs in the United States and Canada." *Evaluation and Program Planning* 20(1):57-75, Schwartz, R. 1998. "The Politics of Evaluation Reconsidered. A Comparative Study of Israeli Programs." *Evaluation* 4(3):294-309.

[140] Taylor, D. 2005. "Governing Through Evidence: Participation and Power in Policy Evaluation." *Journal of Social Policy* 34(04):601-18.

[141] For instance, the evidence-based policymaking model 'advocates a more rational, rigorous and systematic approach, and moves beyond traditional notions of research to adopt a broader

> Evaluations can have a strong influence in the process of justifying a particular policy recommendation or in the competition for jurisdiction over an issue within a democratic institution [...] there must be some shared agreement on the nature of evaluation and how it is to be conducted.[142]

This does not remove values and beliefs from the political discussion, but it uses the power vested in this activity to support the positions of stakeholders.[143] This power can also be perceived in their influence over the evaluation process.[144]

In addition to the variables discussed by Flinders, there are two other elements relevant to the core theme of this thesis. The first one is *efficiency* as a value that underpins policymaking. Reforms to improve the performance of organisations were part of the early issues that contributed to the development of this concept.[145] International organisations promoted reforms to 'make markets work efficiently and corrective interventions where there are market failures.'[146] These reforms recognised the importance of injecting competition into the delivery of services and the incorporation of mechanisms to increase efficiency, i.e. 'one stop window counter' and e-government[147], whose aim was ...

> ... to modify the Weberian bureaucratic model of the state [...] based on a series of principles, including the impersonal nature of working relations, the standardization of working procedures and routines, civil servant recruitment and promotion based on technical and professional progress, rationality in the division of labour and the establishment of authority hierarchies.[148]

understanding.' Sutcliffe, S., and J. Court. 2005. "Evidence-Based Policymaking: What is it? How does it work? What relevance for developing countries?", edited by Overseas Development Institute. P. 1.

[142] Zapico-Goñi, E. 2007. "Matching Public Management, Accountability and Evaluation in Uncertain Contexts: A Practical Suggestion." *Evaluation* 13(4):421-38.

[143] Hanberger, A. 2006. "Evaluation of and for Democracy." Ibid.12(1):17-37.

[144] Cook, J. R. 2014. "Using Evaluation to Effect Social Change: Looking Through a Community Psychology Lens." *American Journal of Evaluation* 36(1):107-17.

[145] Hansson, F., M. T Norn, and T. Bungaard Vad. 2014. "Modernize the public sector through innovation? A challenge for the role of applied social science and evaluation." *Evaluation* 20(2):244-60.

[146] The World Bank. 1992. *Governance and development.* Washington, D.C., US: The World Bank. P. 6.

[147] Osborne, D., and P. Plastrik. 1998. *Banishing Bureaucracy: The Five Strategies for Reinventing Government.* New York, US: Plume.

[148] García López, R. , and M. García Moreno. 2011. "Managing for Development Results. Progress and Challenges in Latin America and the Caribbean." Washington, D.C., US: Inter-American Development Bank. P. 3.

This involved important organisational changes and modifications in terms of politics. The institutionalisation of evaluation is a clear example of the adoption of an instrument focused on the idea of increasing efficiency through the generation of evidence 'to improve policy performance of the governance arrangements in question.'[149] Nonetheless, the modification of the *status quo* may provoke reactions from the different stakeholders involved in policy. As Prince states: 'Reinventing government is about politics as much as administration.'[150] This raises questions about how evaluation fits in this new model in which efficiency is a central element for government, but not necessarily for all stakeholders.

Adaptability is the other element worth discussing. It can be defined as the ability of the state to overcome new challenges in a successful way and the capacity of institutions to adjust to new conditions. This idea sits within what Duit and Galaz labelled as *flexible governance.* For them, this kind of governance 'has well-developed capacities for exploration' as it can adjust its processes and institutions to a new context.[151] In the words of Lobel: 'The new governance model also requires adaptability and constant learning, recognizing the inevitability and fertility of change while treating ambiguity as an opportunity rather than a burden to overcome.'[152] What this reveals is the relevance of rules for the interaction of stakeholders and the capacity of public organisations to internalise new processes and to adopt new routines. It also comprises the political willingness of stakeholders to respect the new set of rules agreed. In this context: 'An evaluation system can help construct a new research governance structure: evaluation becomes one driver of institutional change.'[153] Evaluation needs to operate in a context that allows the recognition of its

---

[149] Peters, B. G. 2012. "Information and Governing: Cybernetic Models of Governance." in *The Oxford Handbook of Governance*, edited by D. Levi-Faur. Oxford, UK: Oxford University Press. P. 116.

[150] Prince, M. J. 2000. "Banishing Bureaucracy or Hatching a Hybrid? The CanadianFood Inspection Agency and the Politics of Reinventing Government." *Governance* 13(2):215-32. P. 228.

[151] Duit, A., and V. Galaz. 2008. "Governance and Complexity—Emerging Issues for Governance Theory." Ibid.21(3):311-35. P. 322.

[152] Lobel, O. 2012. "New Governance as Regulatory Governance." in *The Oxford Handbook of Governance*, edited by D. Levi-Faur. Oxford, UK: Oxford University Press. P. 67.

[153] Molas-Gallart, J. 2012. "Research Governance and the Role of Evaluation: A Comparative Study." *American Journal of Evaluation* 33(4):583-98. P. 585.

legitimacy when stakeholders agree on the validity of the criteria to evaluate and the reliability of findings.

The support to accountability by evaluation and auditing mechanisms, for instance, is limited by the capacity of organisations to modify their processes, bureaucratic routines, and legal frameworks to establish new channels of communication for stakeholders.[154] Another element that illustrates this regards the way in which evaluation fits into the new roles of the state. If, as Flinders points out, the state has moved 'from a provider to a commissioner of public services',[155] it is worth reflecting on how these new functions need to be evaluated when government cannot be labelled as the direct provider of public services and its responsibilities have shifted; criteria for evaluation and accountability need also to be reconsidered.

This discussion about governance and the transformation of the state sheds new light about the conditions that have fostered the use of evaluation. The emergence of new actors in the public sphere demands that new institutional arrangements regulate their interaction, including control and risk-reduction mechanisms. Moreover, the uneven distribution of power and the adoption of new schemes of control have also promoted evaluation as a means to adapt to new models of governing. As an umbrella concept, governance shows the current conditions in which evaluation takes place, both as a managerial and as a political instrument. Regarding the latter, evaluation can represent a means for strengthening democracy 'through specific evaluation approaches that reinforced participation, warranted transparency, promoted public welfare.'[156] The linkage between evaluation and democracy is discussed more carefully later in this chapter. The next section focuses on evaluation within the policy process.

2.1.2 Public Policy Analysis

The attention of social problems through different mechanisms of government

---

[154] See Wright, B. E., R. K. Christensen, and K. Roussin Isett. 2013. "Motivated to Adapt? The Role of Public Service Motivation as Employees Face Organizational Change." *Public Administration Review* 73(5):738-47.

[155] Flinders, M. 2008. "The Future of the State." *The Political Quarterly* 79(S1):19-40. P. 21.

[156] Stame, N. 2006. "Governance, Democracy and Evaluation." *Evaluation* 12(1):7-16. P. 7.

intervention has been systematically studied.[157] Policy analysis emerged as an attempt to comprehend the government's response to these problems. Public policy is seen as 'the product of the activity of an authority invested with public power and governmental legitimacy'[158] or, as Knoepfel et al. define it:

> A series of intentionally coherent decision or activities taken or carried out by different public and sometimes private actors, whose resources, institutional links and interest vary, with a view to resolving in a targeted manner a problem that is politically defined as collective in nature.[159]

It can also be conceptualised as a materialisation of those decisions produced by the political process translated into specific actions and programmes.[160] Most definitions recognise the existence of an issue that affects the public interest and the need for intervention, when government is identified as the responsible actor because it holds legitimate public power.

The analysis of public policy can focus on different areas, e.g. on the relationships established between stakeholders or on the mechanisms to address public problems.[161] The existing approaches are 'attempts at analysing how public policies typically evolve or should evolve.'[162] These present both advantages and disadvantages; however, it is possible to carry out a robust analysis acknowledging and making explicit their limitations.

The policy-cycle approach is useful for studying the role of evaluation in policymaking. It conceptualises policy as a set of stages[163] by looking at the different moments that occur after a political decision is made and before a

---

[157] See: Hill, M. J. 2005. *The Public Policy Process*. London, UK: Routledge, John, P. 1999. *Analysing Public Policy*. London, UK: Continuum, Knoepfel, P., C. Larrue, F. Varone, and M. Hill. 2011. *Public Policy Analysis*. London, UK: Policy Press, Parsons, W. 1995. *Public Policy: An Introduction to the Theory and Practice of Policy Analysis*. London, UK: Edward Elgar Publishing Ltd.

[158] Mény, Y., and J. C. Thoening. 1992. *Las políticas públicas*. Barcelona, España: Ariel. P. 89.

[159] Knoepfel, P., C. Larrue, F. Varone, and M. Hill. 2011. *Public Policy Analysis*. London, UK: Policy Press. P. 24.

[160] Howlett, B. W., and A. G. Lewis. 1984. *Policy Analysis for the Real World*. Oxford, UK: Oxford University Press. John, P. 1999. *Analysing Public Policy*. London, UK: Continuum.

[161] *Appendix B* presents a brief summary of some of the most relevant approaches to the analysis of public policy. It is not exhaustive but it allows identifying the core elements of each one of the approaches.

[162] Knill, C., and J. Tosun. 2012. *Public Policy. A New Introduction*. New York, US: Palgrave Macmillan. P. 7.

[163] Lasswell is a key reference for the study of public policy as a discipline as well as for the study of public policy as a set of interrelated stages. Lasswell, H. D. 1970. "The emerging conception of the policy sciences." *Policy Sciences* 1(1):3-14.

policy outcome is obtained.[164] The policy process is seen as "the expression of the popular will as an 'input' into the political system leading through various processing stages to a policy 'output'"(Figure 2.1).[165]

Figure 2.1. Stages of the Policy Cycle
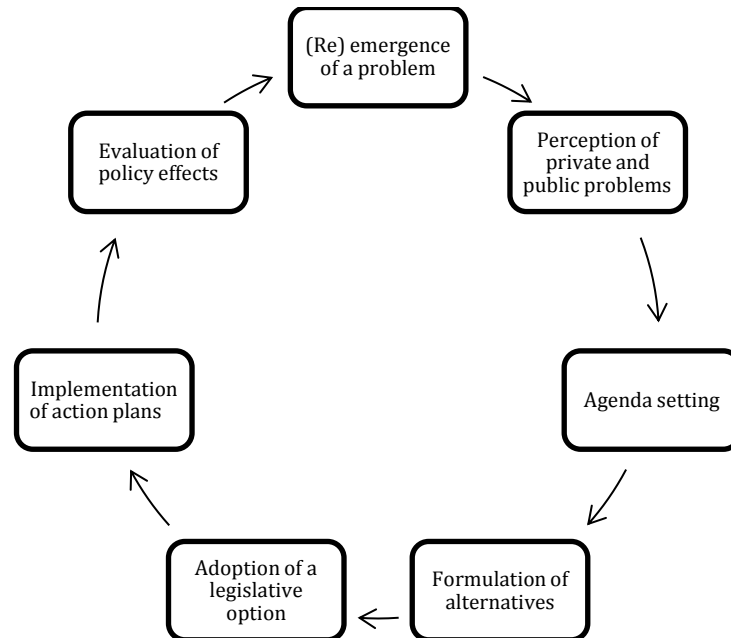
Source: Knoepfel et al., 2011. *Public Policy Analysis.* Policy Press. P. 32.

Despite no consensus existing about the number of stages, most authors agree on stating that the policy cycle starts when a problem gets into the political agenda; followed by the formulation of a possible solution (chosen among different alternatives); then by the implementation of the decision, and concluding with the evaluation of the process.[166] This conceptualisation is not new but it is still relevant for the analysis of the performance of governmental bodies at different levels (local, national, and supranational).[167] It is also useful for grasping the process by disaggregating it into more simplified elements as it 'has proven to provide an excellent heuristic device.'[168] Although as Daniell et al.

---

[164] Hogwood, B. W., and L. A. Gunn. 1984. *Policy Analysis for the Real World*. Oxford, UK: Oxford University Press.

[165] Hill, M. J. 2005. *The Public Policy Process*. London, UK: Routledge. P. 19.

[166] See also: Bardach, E. 2008. *A Practical Guide For Policy Analysis: The Eightfold Path To More Effective Problem Solving* United States of America: Congressional Quarterly, Inc..

[167] See Levy, R. P., M. Barzelay, and A. M. Porras-Gomez. 2011. "The Reform of Financial Management in the European Commission: A Public Management Policy Cycle Case Study." *Public Administration* 89(4):1546-67.

[168] Jann, W., and K. Wegrich. 2007. "Theories of the Policy Cycle." in *Handbook of Public Policy Analysis. Theory, Politics and Methods*, edited by F. Fisher, G. J. Miller, and M. S. Sidney. US: CRC Press. P. 57.

recognise:

> The policy cycle—and rational decision-making processes more generally—are often criticized as rarely reflecting what happens in reality [...] However, the stages noted above still prove useful in practice to orientate needs and thinking around different forms of policy analysis.[169]

Another critic of this approach was Lindblom, who was sceptical about seeing public policy as a linear process as he considered that it implies handling complex social problems.[170] This was also acknowledged by Jann and Wegrich in relation to the distance that stands between this model and the reality:

> While the policy cycle framework takes into account the feedback between different elements of the policy process [...] it still presents simplified and ideal-type model of the policy process [...] these processes do not evolve in a pattern of clear-cut sequences; instead the stages are constantly meshed entangled in an ongoing process.[171]

The limitations of this approach have been recognised.[172] The complexity of looking at policy as a process relies on the fact that stages do not occur sequentially, these are interrelated and may even occur simultaneously.[173] These stages have been studied individually and collectively.[174] For the purposes of this thesis, only a succinct description of each stage is provided, including a brief introduction to evaluation, which although it is conceptualised as the last stage of the process, in practice it is an activity intended to feed back

---

[169] Daniell, K. A., A. Morton, and D. Ríos Insua. 2015. "Policy Analysis and Policy Analytics." *Annals of Operations Research*. P. 3.

[170] Lindblom, C. E. 1959. "The Science of "Muddling Through"." *Public Administration Review* 19(2):79-88.

[171] Jann, W., and K. Wegrich. 2007. "Theories of the Policy Cycle." in *Handbook of Public Policy Analysis. Theory, Politics and Methods*, edited by F. Fisher, G. J. Miller, and M. S. Sidney. US: CRC Press. P. 44.

[172] Parsons, W. 2007. *Políticas Públicas*. Mexico City: Flacso México.

[173] Hill, M. J. 2005. *The Public Policy Process*. London, UK: Routledge.

[174] In terms of agenda-setting see for example: Green-Pedersen, C., and S. Walgrave (Eds.). 2014. *Agenda Setting, Policies, and Political Systems: A Comparative Approach*. Chicago, US: The University of Chicago Press, Kingdon, J. W. 1995. *Agendas, Alternatives and Public Policies*. New York, US: Harper Collins. For the formulation/design stage see: Jordan, A. J., and J. R. Turnpenny (Eds.). 2015. *The Tools of Policy Formulation: Actors, Capacities, Venues and Effects*. UK: Edward Elgar Publishing. The implementation stage can be studied through the work of: Bardach, E. 1977. *The implementation game: What happens after a bill becomes a law*. Massachusetts, US, Mazmanian, D. A., and P. A. Sabatier. 1983. *Implementation and Public Policy*. Illinois, US: University Press of America, Pressman, J. L., and A. B. Wildavsky. 1984. *Implementation: How great expectations in Washington are dashed in Oakland or, why it's amazing that federal programs work at all, this being a saga of the Economic Development Administration as told by two sympathetic observers who seek to build morals on a foundation of ruined hopes*. California, US: University of California Press.

into the whole cycle. Its contribution can be perceived from the fact that evaluation evidence can guide policymakers to making more informed decisions and to reflect on those made. It is a 'powerful tool to reframe an issue once thought to be resolved by policymakers.'[175]

The agenda-setting stage can be considered the initial part of the process. This 'is the list of subjects of problems to which governmental officials are paying some serious attention at any given time.'[176] It involves the acknowledgement of a public problem and the need for a government's intervention. Its connection with the hidden politics of evaluation is that during this stage, negotiation of interests takes place:

> Even though the agenda is not just shaped by group power, but by outside events or by luck, agenda setting models do not say that ideas have an independent existence, merely that a variety of factors conjoin in numerous and diverse ways to shape the political agenda.[177]

Multiple actors with different interests can determine the problems that get into the public agenda and that might influence the process.[178] This involves not only their capacity to promote a particular issue but also to transmit effectively the logic and the values behind it.[179] The characterisation of policy problems will define the nature of the criteria for evaluation and its suitability.[180] Evaluation can provide evidence to support the inclusion (or exclusion) of a particular issue in the agenda,[181] i.e. by encouraging innovation,[182] and also to identify and quantify policy problems.

---

[175] Gerston, L. N. 2010. *Public Policy Making: Process and Principles*. New York, US: M.E. Sharpe. P. 112.

[176] Kingdon, J. W. 1995. *Agendas, Alternatives and Public Policies*. New York, US: Harper Collins. P. 3.

[177] John, P. 1999. *Analysing Public Policy*. London, UK: Continuum. P. 147.

[178] See Weible, C. M., T. Heikkila, P. deLeon, and P. A. Sabatier. 2012. "Understanding and Influencing the Policy Process." *Policy Sciences* 45(1):1-21.

[179] Nelson presents case studies that explore how a policy issue is incorporated into the public agenda and the role of important stakeholders such as politicians and the media. Nelson, B. J. 1984. *Making an Issue of Child Abuse: Political Agenda Setting for Social Problems*. Illinois, US: University of Chicago Press.

[180] House, E. R., and K. R. Howe. 1999. *Values in Evaluation and Social Research*. London, UK: Sage.

[181] Henry, G. T., and M. M. Mark. 2003a. "Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions." *American Journal of Evaluation* 24(3):293-314.

[182] Hansson, F., M. T Norn, and T. Bungaard Vad. 2014. "Modernize the public sector through innovation? A challenge for the role of applied social science and evaluation." *Evaluation* 20(2):244-60.

Policy formulation regards the establishment of alternatives of solution according to the resources available and the expected goals from government's intervention.[183] It 'defines the legal bases for the objective intervention, instruments, and the operational arrangements of the public action.'[184] The core of this stage is decision-making, which can be studied from different perspectives. The bounded rationality approach developed by Simon establishes that the selection of alternatives is based on the maximisation of the decision-maker's values and it takes place after an exhaustive analysis of the alternatives.[185] In contrast, the incremental model proposed by Lindblom suggests that a policy decision has to be made, based on a successive limited comparisons approach to reduce the risk of failure.[186]

Evaluation outputs are expected to provide elements to reduce uncertainty in the decision-making and the risk inherent to the context in which policy occurs.[187] This information can also be used for making diagnostics that strengthen the design of policy instruments.[188] As Sanderson suggests: 'Evaluation is required to assess and understand how policies have worked (or not) and why, so that lessons can be learned to inform improvements.'[189] This will help reducing the uncertainty that stakeholders face and adjust their decisions in the future.

---

[183] Hill, M. J. 2005. *The Public Policy Process*. London, UK: Routledge.
[184] Knoepfel, P., C. Larrue, F. Varone, and M. Hill. 2011. *Public Policy Analysis.* London, UK: Policy Press. P. 157.
[185] Simon, H. A. 1957. *Administrative Behavior: A study of decision-making processes in administrative organisation*. New York, US: Macmillan.
[186] Lindblom, C. E. 1959. "The Science of "Muddling Through"." *Public Administration Review* 19(2):79-88.
[187] The utility of evaluation and its role as an instrument directed to inform policymaking is one of the most relevant discussions in the evaluation studies literature (see Section 2.2.2). It depends on multiple variables out of the evaluator's control and important efforts need to be made to foster the effective use of evaluation in the public sector. See: Feinstein, O. N. 2002. "Use of Evaluations and the Evaluation of their Use." *Evaluation* 8(4):433-39, Johnson, K., L. O. Greenseid, S. A. Toal, J. A. King, F. Lawrenz, and B. Volkov. 2009. "Research on Evaluation Use. A Review of the Empirical Literature From 1986 to 2005." *American Journal of Evaluation* 30(3):377-410, Leviton, L. C., and E. F. X. Hughes. 1981. "Research On the Utilization of Evaluations: A Review and Synthesis." *Evaluation Review* 5(4):525-48.
[188] For example, in Mexico, CONEVAL presents diagnostic information about policy areas like health, education, housing and social security through the publication of the *Social Policy Evaluation Report.*
[189] Sanderson, I. 2002. "Evaluation, Policy Learning and Evidence-Based Policy Making." *Public Administration* 80(1):1-22. P. 19.

Policy implementation,[190] one of the most studied stages of the policy cycle,[191] refers to…

> … the carrying out of a basic policy decision […] Ideally, that decision identifies the problem(s) to be addressed, the objective(s) to be pursued and in a variety of ways "structures" the implementation process.[192]

Sabatier developed the top-down and bottom-up approaches for grasping this stage. The top-down approach 'starts with a policy decision and focuses on the extent to which its objectives are attained over time and why.'[193] On the opposite, the bottom-up approach …

> … starts by defining the network of actors involved in service delivery in one or more local areas and asks them about their goals, strategies, activities and contracts […] then it uses the contacts as a vehicle for developing a network technique to identify the local, regional and national actors involved in the planning, financing and execution of the relevant governmental and non-governmental programs.[194]

This gives policy analysts different lenses for understanding how policy is transformed into outputs. From the systemic perspective, it is possible to look at this as a relationship between inputs and outputs. However, according to Birkland, this perspective does not explain clearly how these outputs are transformed, because it 'treats the political system as a *black box.*'[195] He highlights the importance of dissecting what happens during the execution of political decisions. The principal-agent model is a useful tool for comprehending

---

[190] The work of Pressman and Wildavsky is a key reference for the analysis of the main challenges that policymakers face during implementation, particularly about how a specific policy can easily fail due to a number of factors that were not (or could not be) considered during the formulation stage. Pressman, J. L., and A. B. Wildavsky. 1984. *Implementation: How great expectations in Washington are dashed in Oakland or, why it's amazing that federal programs work at all, this being a saga of the Economic Development Administration as told by two sympathetic observers who seek to build morals on a foundation of ruined hopes*. California, US: University of California Press.

[191] See Barrett, S. M., and C. Fudge (Eds.). 1981. *Policy and Action: Essays on the Implementation of Public Policy*. London, UK: Methuen, Mazmanian, D. A., and P. A. Sabatier. 1983. *Implementation and Public Policy*. Illinois, US: University Press of America, Palumbo, D. J., and M. A. Harder (Eds.). 1981. *Implementing Public Policy*. Massachusetts, US: Lexington Books.

[192] Mazmanian, D. A., and P. A. Sabatier. 1983. *Implementation and Public Policy*. Illinois, US: University Press of America. P. 20.

[193] Sabatier, P. A. 1997. "Top-down and bottom-up approaches to implementation research." in *The Policy Process. A Reader*, edited by M. Hill. London, UK: Pearson Prentice Hall. P. 281.

[194] Ibid. P. 281.

[195] Birkland, T. A. 2011. *An Introduction to the Policy Process: Theories, Concepts, and Models of Public Policy Making*. New York, US: M.E. Sharpe.

the relationship between policy-designers and operators.[196] This model explores the interaction between two individuals (the principal and the agent) in which the principal …

> … would like to delegate a task to the agent, he wants to determine what (minimal) reward he should offer the agent to ensure that 1) the agent accepts to perform the task, and 2) the agent performs the task in a satisfactory way.[197]

The discretionary behaviour of street-level bureaucrats[198] responds to the challenges that they face when implementing and which existence was not or could not have been considered during formulation.[199] This stage represents an area of the process where control may be diluted and that might affect how incentives work in the public sector.[200] There, evaluation can analyse the activities and processes that take place during implementation. The dichotomy between the top-down and bottom-up approaches shows that the main challenge for policymakers is overcoming the potential difficulties that emerge during implementation and strengthening the communication and information channels between them and street-level bureaucrats.[201] Evaluation can give principals more elements for coping with the discretionary behaviour of agents, reducing for them the risk of distorting objectives. Evaluation can also dissect the operation of a programme to identify those elements that can be improved.[202] This finds echoes in the discussions about efficiency, as evaluation findings can expose unnecessary processes with which important savings can

---

[196] This resonates with the concept of 'street-level bureaucrats' which refers to the influence that is exercised by those public officers in charge of the execution of a decision taken at the top of the organization and how this affects the policy output. See: Peters, B. G., and J. Pierre. 2012. *The SAGE Handbook of Public Administration*. Los Angeles, US: SAGE., Ellis, K. 2011. "'Street-level Bureaucracy' Revisited: The Changing Face of Frontline Discretion in Adult Social Care in England." *Social Policy & Administration* 45(3):221-44, Lipsky, M. 1971. "Street-Level Bureaucracy and the Analysis of Urban Reform." *Urban Affairs Review* 6(4):391-409.

[197] Van Ackere, A. 1993. "The principal/agent paradigm: Its relevance to various functional fields." *European Journal of Operational Research* 70(1):83-103. P. 83.

[198] Laws, D., and M. Hajer. 2006. "Policy in Practice." in *The Oxford Handbook of Public Policy*, edited by M. Moran, M. Rein, and R. E. Goodin. Oxford, UK: Oxford University Press.

[199] Hill, M. J. (Ed.). 1997. *The Policy Process: A reader*. London, UK: Routledge.

[200] Dixit, A. 2002. "Incentives and Organizations in the Public Sector: An Interpretative Review " *Journal of Human Resources* 37(4):696-727.

[201] Bache, I., L. Reardon, I. Bartle, M. Flinders, and G. Marsden. 2014. "Symbolic Meta-Policy: (Not) Tackling Climate Change in the Transport Sector." *Political Studies* 63(4):830-51.

[202] Lion, C., P. Martini, and S. Volpi. 2006. "Evaluating the Implementation Process: A Contribution within the Framework of the European Social Fund (ESF) Programme." *Evaluation* 12(3):313-29.

be made. Politicians for example, would have more arguments to legitimise decisions about budget cuts. This also resonates with the idea of the 'black box' about the need to identify and analyse all substantive activities that need to be performed to achieve the expected result.[203] The identification of key problematic procedural areas is perhaps one of the most substantial contributions of evaluation to implementation, and knowing why policies succeed or fail is a permanent concern for policymakers;[204] moreover, understanding how stakeholders interact during implementation matters in terms of the democratic credibility of the process.[205]

As the core theme of this thesis, evaluation is discussed extensively later in this chapter. Only a succinct overview is presented to introduce this concept. Evaluation 'represents a mechanism to identify policy results, its justification and its effectiveness.'[206] It can be defined as …

> … the systematic and objective assessment of an on-going or completed project, program, or policy, including its design, implementation, and results […] to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact, and sustainability.[207]

A central element of these definitions is the capacity of evaluation to obtain reliable information about different aspects of policy, although evaluation can serve multiple purposes.[208] Reviewing the policy process allows identifying the contribution of evaluation in terms of its capacity to measure outcomes that can

---

[203] An interesting example is the work of Alexander who studied the use of evaluation in the National Health Service (NHS) by describing 'apparent differences in the impact of three health-service evaluations.' In her study, Alexander highlights the relevance of the people involved in the evaluation and how this affects the implementation of recommendations. Alexander, H. 2003. "Health-Service Evaluations: Should We Expect the Results to Change Practice?" Ibid.9(4):405-14.

[204] Barrett, S. M. 2004. "Implementation Studies: Time for a Revival? Personal Reflections on 20 Years of Implementation Studies." *Public Administration* 82(2):249-62. O'Toole, L. J. Jr. 2000. "Research on Policy Implementation: Assessment and Prospects." *Journal of Public Administration Research and Theory* 10(2):263-88.

[205] Bache, I., and R. Chapman. 2008. "Democracy through Multilevel Governance? The Implementation of the Structural Funds in South Yorkshire." *Governance: An International Journal of Policy, Administration and Institutions* 21(3):397-418.

[206] Guerrero, J. P. 1995. "La evaluación de las políticas públicas: enfoques teóricos y realidades en nueve países desarrollados." *Gestión y Política Pública* IV(1):47-115. P. 49.

[207] Zall Kusek, J., and R. C. Rist. 2004. "Ten Steps to a Results-Based Monitoring and Evaluation System." edited by The World Bank. Washington, D.C., US. P. 12.

[208] See Segone, M. (Ed.). 2008. *Bridging the gap. The role of monitoring and evaluation in evidence-based policy making*: UNICEF.

provide evidence about how governments deliver.[209] Also, it has been useful for looking at the specific contributions of evaluation to the different stages of the process.[210]

Evaluation transcends the task of verifying the accomplishment of goals or objectives. It produces important amounts of evidence directed to different purposes, some of them political.[211] Evaluation can represent an important asset for stakeholders when policy ends are in conflict, e.g. efficiency vs. equity.[212]   In the light of governance, the policy process, as the translation of political decisions into action, is now constrained by the interests of multiple stakeholders. Consequently, risk-reduction and control mechanisms acquire more relevance. In terms of control, it is useful to deepen the relationship between evaluation and performance management, developed in the next section.

2.1.3 Performance Management

Performance-management systems have introduced monitoring tools for generating information about different aspects of public policy, particularly through the use of indicators.[213] A monitoring and evaluation system …

> … provides government officials, development managers, and civil society with better means for learning from past experience, improving service delivery, planning and allocating resources, and demonstrating results as part of accountability to key stakeholders.[214]

In conjunction with evaluation, performance measurement[215] is one of the

---

[209] Rotberg, R. I. 2014. "Good Governance Means Performance and Results." *Governance* 27(3):511-18.

[210] See Gris Legorreta, P. C. 2011. "La evaluación de políticas públicas como instrumento para la rendición de cuentas." *Trimestre Fiscal* 99:89-103.

[211] Saunders, M. 2012. "The use and usability of evaluation outputs: A social practice approach." *Evaluation* 18(4):421-36.

[212] Winship, C. 2008. "Policy Analysis as Puzzle Solving." in *The Oxford Handbook of Public Policy*. New York: Oxford University Press.

[213] According to the WB, a performance management system has as its primary objectives 'to learn, to steer and control and to provide accountability.' Dorotinsky, W. 2011. *Implementing Results Based Budgeting in Thailand*. Washington D.C., US: The World Bank. P. 6.

[214] The World Bank. 2004. *Monitoring & Evaluation: Some Tools, Methods & Approaches*. Washington, D.C. P. 5.

[215] Frequently, the terms 'performance management' and 'performance measurement' are used indistinctively. This is a common misinterpretation, performance management refers to the set of tools and instruments (e.g. human resources systems, technology, etc.) used to the accomplishment of organisational goals, performance measurement is considered a tool of

pillars of the NPM directed to assess public policy and to identify goals.[216] Both activities sit within the performance-management model that overviews governmental activity.[217] Nonetheless, it is important to distinguish these two activities as they aim at different purposes and are implemented through different methods (Table 2.1).

Table 2.1. Main Differences between Performance Monitoring and Evaluation

	Performance monitoring	Evaluation
Focus	On organisations	On policies and programmes
Time	Periodic	Episodic
Methods	Mainly quantitative	Mixture of qualitative and quantitative
Who	Mostly conducted or commissioned by Finance, Operations, or even Personnel branches	Mostly conducted or commissioned by policy branches

Source: Talbot, C.,2010 *Performance in Government. The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom*. The World Bank. P. 12.

The differences observed by Talbot serve the purpose of deepening into the distinction between outputs and outcomes.[218] While monitoring is supposed to focus on the efficiency of public services (outputs), evaluation concentrates on the identification of the effects (outcomes). This distinction can be observed through the example of the *Cold Weather Plan for England* whose purpose was 'to enhance resilience in the event of severe cold weather.'[219] This implies the execution of different activities undertaken by health and social-care agencies that interact with risk groups of the population who are more vulnerable to the effects of cold weather. The monitoring of this plan could focus on measuring the number of influenza vaccinations provided to prevent the risk of disease. This establishes a connection between inputs and a quantifiable output. In terms of outcomes, the interventions directed to foster the participation of

---

performance management.

[216] Talbot, C. 2010h. *Theories of Performance: Organizational and Service Improvement in the Public Domain*. Oxford, UK: Oxford University Press.

[217] Hunter, D. E. K., and S. B. Nielsen. 2013. "Performance Management and Evaluation: Exploring Complementarities." *New Directions for Evaluation* 2013(137):7-17.

[218] Talbot, C. 2010a. "Performance in Government. The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom." Washington D.C., US: The World Bank.

[219] Public Health England. 2014b. "Cold weather plan for England 2014." London, UK.

women in the labour market are useful for illustrating the role of evaluation. A long-term effect (outcome) could be women's financial independence, which is a much more complex variable to quantify and that might imply a subjective judgement.[220]

Performance measurement fits into the idea of control that takes place in the public sector.[221] Van Dooren et al. recognise five major stages of performance measurement (see Figure 2.2 below). These are consistent with the five elements observed by Pollitt (activity, measurement, data, criteria, and use) that manifest its logic as an instrument for the generation of evidence that can feed into other systems, i.e. evaluation and accountability.[222] Evaluators, for example, can use indicators as an aid for making their judgements about a policy, regardless of engaging in more complex studies for analysing policy.[223]

Figure 2.2 An Ideal-typical Model of the Performance Measurement Process
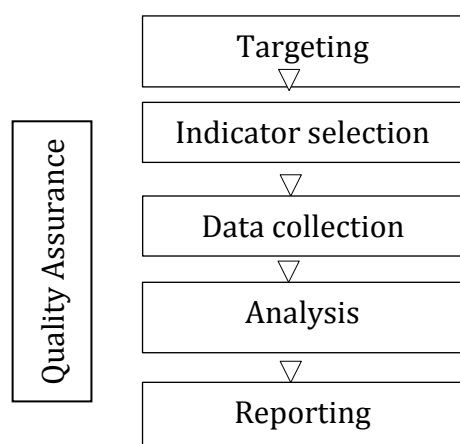
Source: van Dooren et. al, 2010. *Performance Management in the Public Sector*. Routledge. P. 55.

Taking into consideration the ideas about 'post-bureaucratic organisations'[224], the focus on the identification of outcomes instead of processes gives evaluation

---

[220] An additional example of measuring outputs and outcomes is provided by the analysis of the "Stop AIDS Love Life" campaign in Ghana carried out by the John Hopkins University. From this experience, it was identified as an output to be measured 'the number of HIV tests performed' which is a result derived from an activity; while as an outcome they identified 'health behaviour' as a consequence to the exposure to the campaign and the knowledge transmitted through the different media (pamphlets, TV, radio, etc.). The John Hopkins University. 2006. "Monitoring Outputs and Outcomes And Introduction to Study Design."

[221] Behn, R. D. 2003. "Why Measure Performance? Different Purposes Require Different Measures." *Public Administration Review* 63(5):586-606.

[222] Pollitt, C. 2013. "The logics of performance management." *Evaluation* 19(4):346-63.

[223] Hunter, D. E. K., and S. B. Nielsen. 2013. "Performance Management and Evaluation: Exploring Complementarities." *New Directions for Evaluation* 2013(137):7-17.

[224] As Kernaghan states, this kind of organisations adopt new models related to a set of

and auditing more leverage, while it also imposes the need for developing instruments for organisations to make explicit its contribution to the public sector. This is also underpinned by the idea that organisations need to be 'fuelled by influence rather than power.'[225] Therefore: *How can this influence be widened?* This question highlights the relevance of building evaluation and audit systems considered legitimate by stakeholders. It transcends the role of performance management as a mechanism of monitoring public policy, as it involves using that information for making judgements.

Performance monitoring is embedded in a logic of technical and objective use of information, but the politics involved needs to be taken into account for understanding the behaviour of the actors that participate in the production and use of performance data, e.g. to look at how 'politicians may use the formulation of performance measures to make their political priorities explicit, while agents use them to guide their priorities and performance.'[226] Pollitt's ideas about the 'sub-logics' of performance management can help evince some of the political issues involved in this instrument. Some of these 'sub-logics' expose the discretion that actors exercise in the design and use of targets and indicators, e.g. *cheating*—bending or breaking the rules.[227]

Although the implementation of these systems is intended to increase efficiency in the public sector, an ample range of scholars has questioned its utility.[228] This, according to Propper and Wilson, might be associated with some design flaws:

> First, there is some consensus that gross outcomes or levels based on
> PM [performance measures] do not provide a sufficiently accurate

characteristics important to mention: results-oriented, citizen-centred, competitive, non-departmental and decentralized structure, among others. Kernaghan, K. 2000. "The Post-Bureaucratic Organization and Public Service Values." *International Review of Administrative Sciences* 66(1):91-104.

[225] Finlayson, A. 2012. "Cameron, Culture and the Creative Class: The Big Society and the Post-Bureaucratic Age." *The Political Quarterly* 82(S1):35-47. P. 42.

[226] Bjørnholt, B., and F. Larsen. 2014. "The politics of performance measurement: 'Evaluation use as mediator for politics'." *Evaluation* 20(4):400-11. P. 404.

[227] Pollitt, C. 2013. "The logics of performance management." Ibid.19:346-63.

[228] See: McLean, I., D. Haubrich, and R. Gutiérrez-Romero. 2007. "The Perils and Pitfalls of Performance Measurement: The CPA Regime for Local Authorities in England." *Public Money & Management* 27(2):111-18, Moynihan, D. P. 2006. "Managing for Results in State Government: Evaluating a Decade of Reform." *Public Administration Review* 66(1):77-89.

picture of the relative performance of public sector organisations…
Second, a single PM is not sufficient. Public sector organisations often have multiple stakeholders who have differing, and sometimes conflicting, goals…
Third, the intended purpose for each potential measure should dictate both its form and the decision whether to publicly disclose the resulting performance information.[229]

The core question that arises here is: *Which can be considered the reliable mechanisms to assess an organisation's performance in relation to the political power that it holds?* This means: *To what extent can self-imposed indicators reflect accurately the achievements made by an organisation?* And more importantly: *How does this information feed into evaluation processes for judging a policy?* As an input for the development of evaluations, performance-management information needs to be taken into account, considering these political nuances.[230]

This overview of performance measurement is helpful for looking at its relationship with evaluation and how these activities are involved in the politics of the public sector.[231] This first part of the theoretical framework highlights the current conditions in which evaluation takes place (governance), the contribution of this activity to policymaking, and the complementarity that characterises the relationship with performance monitoring. These discussions are relevant for understanding at the macro-level of the relationship between politics and evaluation. For a specific level of analysis, the next section is devoted to the key discussions about evaluation and its institutionalisation.

---

[229] Propper, C., and D. Wilson. 2003. "The Use and Usefulness of Performance Measures in the Public Sector." *CMPO Working Paper Series* 03(073). P. 19.
[230] In Mexico, for example, the use of performance measurement indicators is a core element of some mandatory evaluations (i.e. SPE).
[231] Posner, P. L., and R. Schwarzt. 2011. "Accountability Institutions and Information in the Policy-Making Process." in *Accountable Governance: Problems and Promises*, edited by M. J. Dubnick and H. G. Frederickson. London, UK: Routledge.

2.2 Evaluation Studies

Understanding the role of evaluation within governance and, more specifically, within the policy process, helps filling the gap existent in the knowledge about how and why the political nature of evaluation matters. This section aims to drill down into the discussion about evaluation as a mechanism that is embedded in a political reality that affects its performance and use. Looking at the hidden politics of evaluation from the perspective of the institutional framework contributes to the knowledge about the conditions that favour or hinder political influence and, more importantly, how the rules established for evaluation shape the relationships among stakeholders.

The theoretical development of evaluation has increased during the last decade, as the necessity to value the effectiveness of the public action has become more recurrent in the public sphere.[232] Although there is vast literature concerning evaluation,[233] about both its theoretical and practical grounds,[234] it is neither necessary nor possible to provide an exhaustive theoretical framework. Instead, this section aims to present the most relevant topics about its state of the art, its utilisation, and the relevance of its institutionalisation.

2.2.1 Policy Evaluation: The State of the Art

It is possible to identify two major contributions of evaluation to the public sector. The first one is its capacity to identify the effects of public policy, e.g. impact evaluation.[235] The second is its role as a mechanism for strengthening policymaking by using the information that it produces to make the necessary

---

[232] Henry, G. T., and M. M. Mark. 2003c. "Toward an Agenda for Research on Evaluation." *New Directions for Evaluation* 2003(97):69-80, Rossi, P. H., and S. R. Wright. 1977. "Evaluation Research: An Assessment of Theory, Practice, and Politics." *Evaluation Review* 1(1):5-52.

[233] See Frederiksen, L. F., F. Hansson, and S. B. Wenneberg. 2003. "The Agora and the Role of Research Evaluation." *Evaluation* 9(2):149-72, Nagel, S. (Ed.). 2002. *The Handbook of Public Policy Evaluation*. California, US: SAGE.

[234] In terms of the theoretical bases of evaluation, it is relevant to refer to the work of Alkin, M. C. 2013. *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences*. California, US: SAGE. The study of evaluation practice can be understood through the work of Schwandt, T. A. 2002. *Evaluation Practice Reconsidered*. New York, US: Peter Lang.

[235] See Morton, M. H. 2009. "Applicability of Impact Evaluation to Cohesion Policy." in *An Agenda for a Reformed Cohesion Policy*, edited by European Commission, The World Bank. 2009. "Making Smart Policy: Using Impact Evaluation for Policy Making. Case Studies on Evaluations that Influenced Policy." Washington, D.C., US.

adjustments.[236] The evaluation studies literature identifies multiple applications for evaluation in the public sector.[237] However, for the purposes of this research, this review focuses on three key themes: the conceptualisation of evaluation, the utilisation of its findings, and its institutionalisation. Different authors have understood evaluation in multiple ways.[238] While some think of it as an evidence-generator mechanism,[239] others focus on its role as a tool for measuring the effectiveness of public policy.[240] Its conceptualisation has changed over time as more empirical evidence has been obtained about its contribution to policymaking.[241]

Like other concepts in the policy-analysis literature, there is no unique definition of evaluation. For Picciotto, evaluation 'consists in collecting relevant evidence, identifying suitable evaluative standards and using methods of analysis that are valid and fair.'[242] This definition puts the generation of evidence and the use of robust mechanisms for its obtaining at the core of the concept, but it also leads to reflecting about who will determine the validity and fairness of these mechanisms. In contrast, Parsons emphasises the objective and systematic analysis of policy effects.[243] Authors like Albaek, instead, have given more weight to objectivity:

[236] Bemelmans-Videc, M. L., R. C. Rist, and E. Vedung. 2003. *Carrots, Sticks & Sermons: Policy Instruments and their Evaluation*. London UK: Transaction Publishers.

[237] Rossi, P. H., and S. R. Wright. 1977. "Evaluation Research: An Assessment of Theory, Practice, and Politics." *Evaluation Review* 1(1):5-52, Shadish, W. R., T. D. Cook, and L. C. Leviton. 1991. *Foundations of Program Evaluation: Theories of Practice*. London, UK: SAGE.

[238] The theoretical development of evaluation has been a gradual process. The work of Alkin, Scriven and Shadish et al. are key references in the field. See Alkin, M. C. 2013. *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences*. California, US: SAGE, Scriven, M. 1991. *Evaluation Thesaurus*. London, UK: SAGE, Shadish, W. R., T. D. Cook, and L. C. Leviton. 1991. *Foundations of Program Evaluation: Theories of Practice*. London, UK: SAGE.

[239] Berriet-Solliec, M., P. Labarthe, and C. Laurent. 2014. "Goals of evaluation and types of evidence." *Evaluation* 20(2):195-213, Dunworth, T., J. Hannaway, J. Holahan, and M. Austin Turner. 2008. "Beyond Ideology, Politics and Guesswork: The Case for Evidence-Based Policy." edited by Urban Institute.

[240]Altschuld, J. W., and H. Y. Zheng. 1995. "Assessing the Effectiveness of Research Organizations: An Examination of Multiple Approaches." *Evaluation Review* 19(2):197-216, Chen, H. T. 2005. *Practical Program Evaluation: Assessing and Improving Planning, Implementation, and Effectiveness*. London, UK: SAGE.

[241] Balthasar, A. 2009. "Institutional Design and Utilization of Evaluation: A Contribution to a Theory of Evaluation Influence Based on Swiss Experience." *Evaluation Review* 33(3):226-56.

[242] Picciotto, R. 2005. "Policy Coherence and Development Evaluation: Issues and Possible Approaches." in *Fostering Development in a Global Economy: A Whole of Government Perspective*, edited by OECD. P. 134.

[243] Parsons, W. 1995. *Public Policy: An Introduction to the Theory and Practice of Policy Analysis*. London, UK: Edward Elgar Publishing Ltd.

> The role of evaluation research in rational policymaking is to deliver factual, objective knowledge (i.e. causal but not value premises for decision making) [...] evaluation research will produce objective and value-free knowledge in so far it is based on quantitative positivist social science research procedures.[244]

As a mechanism for reducing uncertainty in decision-making, evaluation 'involves the determination of the impact of policies on target for the direct and timely use of those responsible for a policy intervention.'[245] From these definitions it is possible to see evaluation ...

> As a tool directed to identify the value or merit of public policy;
> As a mechanism to obtain reliable information that can be used for the decision-making process;
> As an instrument to observe the effects of public intervention;
> As an examination of the achievement of goals set by public programmes.

Evaluation offers different alternatives to explain why public programmes have succeeded or not. Its increasing use in the public sector can be explained by several reasons.[246] The first one is its role as a theory-testing activity[247] that allows analysing the possible deviations between what the theory states about the operation of a programme and the results obtained.[248] A good example of this can be found in social programmes directed to fight malnutrition. Theory states that fortified milk might have a positive effect on reducing children's malnutrition;[249] yet, evaluations have shown that this effect is difficult to measure as other factors might intervene, e.g. genetics, quality of milk. The gap between the assumptions made in the design of a programme and the outputs obtained is not a minor issue. Evaluation allows policymakers to analyse the

---

[244] Albaek, E. 1995. "Policy Evaluation: Design and Utilization." in *Policy Evaluation*, edited by Ray C. Rist. Hants, England: Edward Elgar Publishing Company. P. 7.

[245] Owen, J. M., and P. J. Rogers. 1999. *Program Evaluation. Forms and approaches*. Melbourne, Australia: SAGE. P. 32.

[246] Campbell, D., S. Redman, L. Jorm, M. Cooke, A. Zwi, and L. Rychetnik. 2009. "Increasing the use of evidence in health policy: Practice and views of policy makers and researchers." *Australia and New Zealand Health Policy* 6(1):21.

[247] Leeuw, F. L. 1991. "Policy theories, knowledge utilization, and evaluation." *Knowledge and Policy* 4(3):73-91.

[248] This idea has led to the development of programme theory that 'uses substantive knowledge, as opposed to method proclivities, to guide program evaluations [...] it aspires to update, clarify, simplify, and make more accessible the evolving theory of evaluation practice commonly referred to as theory-driven or theory-based evaluation.' Donaldson, S. I. 2005. "Using program theory–driven evaluation science to crack the Da Vinci code." *New Directions for Evaluation* 2005(106):65-84. P. 65.

[249] Torrejón, C. S., C. Castillo-Durán, E. Hertrampf, and M. Ruz. 2004. "Zinc and Iron Nutrition in Chilean Children Fed Fortified Milk Provided by the Complementary National Food Program." *Nutrition* 20(2):177-80.

conditions that might affect policy and to make decisions based on data. However, the way in which this information is produced is an important element to study.

This reveals three important elements regarding the hidden politics of evaluation. First, that evaluation is linked to ideas of rationality, objectivity, and neutrality, which can provide it with a meaning of a highly technical and apolitical activity. Secondly, that there is a set of actors involved in public policy whose interests and preferences can influence the process. Thirdly, that public policy is not the result of a unique decision; it encompasses different values that need to be taken into consideration when evaluating.[250] Like other stages of the policy process, evaluation is embedded in a political context, its conceptualisation as a technical one with the political values at stake.[251] Therefore, the rationality label given to evaluation can be a mechanism to disguise the politics involved in this activity.

Evaluation looks at the relationship between inputs and outputs which entails a series of theoretical assumptions. According to Rossi and Wright, the first assumption is that policy goals are well defined and set the basic premises for an accurate design of the evaluation. [252] This means that there is a clear understanding of the policy problem and of the best policy solution available.[253] In practice, policymaking is not a straightforward process. It demands the interaction of different actors and the establishment of networks, not always driven by rational purposes.[254]

The second assumption is that evaluation takes place in a political context; so it is exposed to the influence of the stakeholders. Although evaluation was

---

[250] Sanderson, I. 2006. "Complexity, 'practical rationality' and evidence-based policy making." *Policy & Politics* 34(1):115-32.

[251] Palumbo, D. J. 1987. *The Politics of Program Evaluation*. California, US: SAGE.

[252] Rossi, P. H., and S. R. Wright. 1977. "Evaluation Research: An Assessment of Theory, Practice, and Politics." *Evaluation Review* 1(1):5-52.

[253] The work of Hansen is relevant because it aims to analyse the strengths and weaknesses of different evaluation models and at the same time to emphasise the importance of this analysis at the design stage of an evaluation. Foss Hansen, H. 2005. "Choosing Evaluation Models: A Discussion on Evaluation Design." *Evaluation* 11(4):447-62.

[254] Hans Klijn, E., and J. F.M. Koppenjan. 2014. "Accountable Networks." in *The Oxford Handbook of Public Accountability*, edited by M. Bovens, R. E. Goodin, and T. Schillemans. Oxford, UK: Oxford University Press.

developed under the presumption that it would bring more rationality into policymaking, in practice this is difficult to achieve as this process is not entirely rational. The use of evidence can provide policymakers with more elements for the design and implementation of policies, but not every decision is based on the information produced by evaluation.[255] The main argument underpinning this idea is that if policymakers obtain reliable and timely information they will use it to make better decisions, but this does not occur *de facto.* Multiple factors can intervene and it is a discretionary decision for the way in which each stakeholder establishes his/her priorities (including using evaluation).

Consequently, evaluation is expected to provide evidence for decision-making, but *what kind of effects do we expect to observe* and *how can those effects be observed?* The adoption of the evidence-based policymaking approach has been largely studied and analysed.[256] Scholars have also acknowledged the political component that is embedded in this approach:

> The production of evidence for policy making is by its nature a political process. But just as the political process of problem solving cannot be based exclusively or even predominantly on "evidence", neither can the production of evidence be based on "science" alone.[257]

Constraining the discussion of policy effectiveness to technical arguments might limit the range of applications of evaluation research to this process.[258] As a discipline derived from the social sciences, evaluation has developed quantitative and qualitative approaches for assessing different features of policy. Policymakers around the world have access to an important range of tools for the evaluation of their policies.[259] The selected approach will determine the nature of the results obtained,[260] the method to apply, and its

---

[255] Rutter, J. 2012. "Evidence and Evaluation in Policy Making. A problem of supply or demand?", edited by Institute for Government.

[256] See for example: Oliver, K., T. Lorenc, and S. Innvær. 2014. "New directions in evidence-based policy research: A critical analysis of the literature." *Health Research Policy and Systems* 12(1).

[257] The LSE GV314 Group. "Evidence-based politics. Government and the production of policy research." P. 19.

[258] du Toit, A. 2012. "Making Sense of 'Evidence'. Notes on the Discursive Politics of Research and Pro-Poor Policy Making." in *Working Paper 21*. Bellville: PLAAS, UWC.

[259] See The World Bank. 2004. *Monitoring & Evaluation: Some Tools, Methods & Approaches*. Washington, D.C.

[260] The literature about evaluation models is vast, as it is not the main focus of this research for further knowledge on this topic it is advisable to refer to the work of Stufflebeam who

impact on the robustness and legitimacy of evaluation.[261] These two factors are central for the political acceptance of an evaluation. The robustness of the methods is crucial for sustaining the findings and recommendations of an evaluation. As Chelimsky discusses:

> Stronger methods allow us more confidence in the value judgments we make about a particular intervention, and weaker ones (which may be the only alternatives feasible in a particular design situation) force us to pepper those judgments with needed caveats.[262]

From an overall perspective, the evaluation process comprises a series of technical and political decisions constrained by the interests of the stakeholders involved. This relates back to the secondary questions of this research about the kind of instruments existent for evaluation and the values behind it. The utility of evaluation is still a central element of the body of literature in evaluation, as it is addressed in the next section.

2.2.2 The Utilisation of Evaluation

Utilisation is perhaps the most analysed topic regarding evaluation.[263] The allocation of resources for evaluation is expected to generate benefits in return. However, the mere performance of an evaluation does not mean per se that outputs will be used, but *what does 'use' mean?*

> Use is generally understood to refer to a direct action that has occurred as a result of an evaluation (i.e. instrumental use) or to something that is newly learned about a programme, its participants, its operations, or outcomes through an evaluation (i.e. conceptual use); the action or learning can take place as a result of evaluation findings, or as a result of participation in evaluation procedures (i.e. process use) [...] in some cases, the concept of

developed a taxonomy of the most relevant evaluation models applied for programme evaluation. See: Stufflebeam, D. 2001. "Evaluation Models." *New Directions for Evaluation* 2001(89):7-98.

[261] The robustness of the evaluation model is linked to the extensive discussion in the literature about experimental and quasi-experimental models for evaluation. The work of Campbell and Stanley explores relevant concepts for the evaluation studies field such as external, internal, and construct validity of this kind of approaches. Campbell, D. T., and J. C. Stanley. 1966. *Experimental and Quasi-Experimental Designs for Research*. Chicago, US: Houghton Mifflin.

[262] Chelimsky, E. 2012. "Valuing, Evaluation Methods, and the Politicization of the Evaluation Process." *New Directions for Evaluation* 2012(133):77-83. P. 78

[263] See: Ginsburg, A., and N. Rhett. 2003. "Building a Better Body of Evidence: New Opportunities to Strengthen Evaluation Utilization." *American Journal of Evaluation* 24(4):489-98, Leviton, L. C., and E. F. X. Hughes. 1981. "Research On the Utilization of Evaluations: A Review and Synthesis." *Evaluation Review* 5(4):525-48, Nelson, L. 2008. "Utilization-focused Evaluation of the Program Evaluation Process of a Missouri School District." Missouri, US: University of Missouri.

use is used to refer to waving the flag of evaluation to claim a rational basis for action (or inaction), or to justify pre-existing positions (i.e. symbolic use).[264]

Recent literature about utilisation has attempted to reach a more specific level of identification of its potential uses (see Figure 2.3 below).[265] The output of an evaluation becomes an input for policymaking and, as discussed, it can represent an important asset for supporting decisions in multiple policy areas.[266] It is also an instrument for discussion in the partisan political arena because the evidence derived from robust research processes can inform decisions about the creation, conservation, or termination of programmes.[267] As Roche and Kelly suggest: 'The most immediate challenge for programs focused on working politically is therefore to demonstrate during the program life that they are leading to tangible outcomes and benefits for people.'[268] This underlines the legitimising function of evaluation. Governments need to demonstrate that their policies are benefiting the population and that resources are being well spent. This makes explicit the connection between evaluation and auditing, as well as between evaluation and accountability. These are complementary activities that rely on the technical capacities of each instrument to generate evidence.[269]

---

[264] Henry, G. T., and M. M. Mark. 2003a. "Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions." *American Journal of Evaluation* 24(3):293-314. P. 294.

[265] See Boyer, J. F., and L. I. Langbein. 1991. "Factors Influencing the Use of Health Evaluation Research in Congress." *Evaluation Review* 15(5):507-32, Feinstein, O. N. 2002. "Use of Evaluations and the Evaluation of their Use." *Evaluation* 8(4):433-39, Saunders, M. 2012. "The use and usability of evaluation outputs: A social practice approach." Ibid.18:421-36.

[266] See for example the work of Gudmundsson, H. 2003. "The Policy Use of Environmental Indicators - Learning from Evaluation Research." *The Journal of Transdisciplinary Environmental Studies* 2(2):1-12.

[267] Molas-Gallart, J. 2012. "Research Governance and the Role of Evaluation: A Comparative Study." *American Journal of Evaluation* 33(4):583-98.

[268] Roche, C., and L. Kelly. 2012. "The Evaluation of Politics and the Politics of Evaluation." Developmental Leadership Program. P. 6.

[269] As Rist states, 'the strength of evaluation lies in not being restricted to the logic and practice of normative analysis, but in being able to use a variety of other methodological approaches as well.' Rist, R. C. 1995. "Management Accountability: The Signals sent by Auditing and Evaluation." in *Policy Evaluation*, edited by R. C. Rist. Hants, England: Edward Elgar Publishing Company. P. 383.

Figure 2.3 Potential Uses of Evaluation

1. •Instrumental use *[Recommendations derived from evaluation are implemented]*

2. •Enlightenment use *[The information provided by the evaluation eventually has an influence in the decision makers over time]*

3. •Process use *[Evaluation has a research function directed to promote organizational learning by stakeholders]*

4. •Symbolic use *[Results are interpreted mostly for political reasons]*

Source: Based on Vanlandingham, G. 2011. "Escaping the Dusty Shelf: Legislative Evaluation Offices' Efforts to Promote Utilization." *American Journal of Evaluation* 32(1).

Empirical evidence has shown that the utilisation of evaluation results might be affected by different variables other than the recognition of its potential benefits.[270] Leviton and Hughes, for example, highlight:

> The relevance of evaluation to the needs of potential users;
> The extent of communication between potential users and producers of evaluation;
> The translation of evaluations into their implications for policy and programmes;
> The credibility or trust placed in evaluations;
> The commitment or advocacy of individual users.[271]

In general, 'The relevance, quality and credibility of evaluation are three of the most important factors explaining evaluation use [in addition to] evaluation implementation and policy setting.'[272] These factors relate to the generation of data and information about a policy at a convenient time and expressed in terms that can be easily understood by decision-makers or users. The robustness and effectiveness of the technical instruments used for evaluation contribute to the

---

[270] See Iriti, J. E., W. E. Bickel, and C. Awsumb Nelson. 2005. "Using Recommendations in Evaluation." *American Journal of Evaluation* 26(4):464-79, Lawrenz, F., A. Gullickson, and S. Toal. 2007. "Dissemination. Handmaiden to Evaluation Use." Ibid.28(3):275-89, McDavid, J. C., and I. Huse. 2012. "Legislator Uses of Public Performance Reports." Ibid.33(1):7-25, Weiss, C. H. 1998b. "Have We Learned Anything New About the Use of Evaluation?" Ibid.19:21-33.

[271] Leviton, L. C., and E. F. X. Hughes. 1981. "Research On the Utilization of Evaluations: A Review and Synthesis." *Evaluation Review* 5(4):525-48. P. 525.

[272] Hanberger, A. 2011. "The real function of evaluation and response systems." P. 330.

credibility of the results.[273] This inevitably leads to reflecting on the political support given to evaluation, both to its process and outputs. Contandriopoulos and Brousselle point out two core dimensions of evaluation that concern its political nature: *ideology and polarization. Ideology* refers to the linkage between an evaluation output and the user's preferences.[274] If there is no connection between these elements, information is more likely to be dismissed or misused. The political risks associated with deciding on the basis of evidence is another element related to the ideology of stakeholders.[275] Involving stakeholders in the process of evaluation seems to be an important element for aligning the user's beliefs with the findings.[276] Therefore, establishing common-beliefs systems could promote not only the utilisation of findings, but also the systematic performance of evaluations and its inclusion in organisational routines.

Polarization refers to the divergent interests that exist in the public sphere that can lead to political confrontation.[277] For example, an evaluation output could simultaneously find approval from one group and rejection from another. The level of polarization affects utilisation and might be explained by the political sensitivity around a policy issue. The ideas of Schwarz and Struhkamp about the relationship between evaluation and trust are pertinent for understanding this more clearly. For them: 'Trust is a mechanism to reduce complexity [...] is used to come to decisions despite incomplete information.'[278] Evaluation findings are used to support decisions in a policy area and to face criticism from political

---

[273] Credibility can be understood as 'a judgement by others [...] that the evaluation is both competent and objective.' Chelimsky, E. 1995. "The Political Environment of Evaluation and What it Means for the Development of the Field: Evaluation for a New Century: A Global Perspective." *American Journal of Evaluation* 16(3):215-25. P. 219. See also: Heilman, J. G. 1983. "Beyond the Technical and Bureaucratic Theories of Utilization: Some Thoughts on Synthesizing Reviews and the Knowledge Base of the Evaluation Profession " *Evaluation Review* 7(6):707-28.
[274] Contandriopoulos, D., and A. Brousselle. 2012. "Evaluation models and evaluation use." *Evaluation* 18(1):61-77.
[275] Rutter, J. 2012. "Evidence and Evaluation in Policy Making. A problem of supply or demand?", edited by Institute for Government.
[276] Campbell, D., S. Redman, L. Jorm, M. Cooke, A. Zwi, and L. Rychetnik. 2009. "Increasing the use of evidence in health policy: Practice and views of policy makers and researchers." *Australia and New Zealand Health Policy* 6(1):21.
[277] Contandriopoulos, D., and A. Brousselle. 2012. "Evaluation models and evaluation use." *Evaluation* 18(1):61-77.
[278] Schwarz, C., and G. Struhkamp. 2007. "Does Evaluation Build or Destroy Trust?: Insights from Case Studies on Evaluation in Higher Education Reform." Ibid.13(3):323-39.

adversaries. Consequently, trusting the process is crucial for this purpose. The central idea is to reduce polarization to promote the acceptance of evaluation findings and the legitimisation of an evaluation system.[279]

Identifying those key factors for the use of evaluation is not an easy task and it requires shaping evaluation results according to the characteristics of the potential users, i.e. policymakers, politicians, media, etc.[280] As evaluation can drive changes in an organisation, there will be interests to cope with. The pertinence of the information and the credibility of the results are considered fundamental for utilisation. Credibility is a necessary condition for its legitimacy, particularly in the political context in which evaluation takes place, when it is necessary to establish a relationship between evaluators and evaluands to foster utilisation.[281]

Values like independence and objectivity are supposed to bring more rationality into policymaking through the use of robust and 'neutral' evidence. The utilisation of evaluation results is based on the reliability of evidence and preserving this quality is imperative. To expect a total separation between evaluation and politics seems naïve. Evaluators are constantly facing pressures during their tasks, and protecting their independence can signify an important challenge, not always guaranteed by the institutional framework.[282] Independence generates the adequate conditions for the design of an evaluation, for the collection and analysis of data, and for the reporting and

---

[279] According to Balthasar, the institutional design plays an important role in the utilization of evaluation findings and it can serve as a mechanism for the reduction of polarization. See Balthasar, A. 2009. "Institutional Design and Utilization of Evaluation: A Contribution to a Theory of Evaluation Influence Based on Swiss Experience." *Evaluation Review* 33(3):226-56.

[280] In relation to this, different studies have been carried out (mostly in the United States) to observe how evaluation results seem more susceptible to be used by decision makers. Brandon, P. R., and J. M. Singh. 2009. "The Strength of the Methodological Warrants for the Findings of Research on Program Evaluation Use." *American Journal of Evaluation* 30(2):123-57, Fleischer, D. N., and C. A. Christie. Ibid."Evaluation Use. Results From a Survey of U.S. American Evaluation Association Members."158-75, Johnson, K., L. O. Greenseid, S. A. Toal, J. A. King, F. Lawrenz, and B. Volkov. Ibid."Research on Evaluation Use. A Review of the Empirical Literature From 1986 to 2005." (3):377-410, Ledermann, S. 2011. "Exploring the Necessary Conditions for Evaluation Use in Program Change." Ibid.

[281] Lonsdale studied the relationship between the NAO and third parties in the context of the audit work carried out by this body. Lonsdale, J. 2008. "Balancing Independence and Responsiveness: A Practitioner Perspective on the Relationships Shaping Performance Audit." *Evaluation* 14(2):227-48.

[282] Picciotto, R. 2011. "The logic of evaluation professionalism." Ibid.17:165-80.

68

dissemination of findings 'in a balanced, fair, and faithful manner.'[283] This struggle for independence has fostered the debate about the advocacy role of evaluators.[284] These challenges imply dealing continuously with the interests of stakeholders and their power to influence the process.[285] Isolating evaluation from those factors seems unlikely. Nonetheless, the inclusion of relevant stakeholders in the process, the use of a robust methodology, and promoting an appropriate dissemination of the findings, i.e. clarity on the scope and limitations, and timing, are strategies that can foster the utility of evaluation.[286]

These issues are usually addressed by the institutional framework and can show the extent to which these rules establish the necessary conditions for the effective conduction of evaluations. Its objective is giving evaluation more formality and generating safeguards for the potential manipulation of the process. The question that arises here is: *To what extent can formal rules and procedures eliminate discretion?* As Hogwood and Gunn state, most of the political sensitivity of evaluation is related to the fact that 'the co-operation (rather than simply compliance) of public officials and clients is often required.'[287] Consequently, the establishment of rules and procedures might incentivise compliance, but this is not enough, as evaluation needs to be internalised as a managerial and political practice. The institutional framework can formally foster—or even impose—the use of evaluation. However, the actual effect of this use in policymaking requires transcending the managerial sphere and transmitting the political power that it holds for stakeholders.

The theoretical discussions presented in this section set the basis for understanding why politics matters for the utilisation of evaluation. Its political nature is manifested in the level of utilisation of findings, in the inclusion of

---

[283] Datta, L. E. 2000. "Seriously Seeking Fairness: Strategies for Crafting Non-partisan Evaluations in a Partisan World " *American Journal of Evaluation* 21(1):1-14. P. 2.

[284] Greene, J. C. 1997. "Evaluation as Advocacy." Ibid.18:25-35, Grob, G. F. 2014. "How to Become an Effective Advocate Without Selling Your Soul." Ibid.35(3):391-97, House, E. R., and K. R. Howe. 1998. "The Issue of Advocacy in Evaluations." *The American Journal of Evaluation* 19(2):233-36.

[285] See Datta, L. E. 1999. "The ethics of evaluation neutrality and advocacy." *New Directions for Evaluation* 1999(82):77-88.

[286] —. 2000. "Seriously Seeking Fairness: Strategies for Crafting Non-partisan Evaluations in a Partisan World " *American Journal of Evaluation* 21(1):1-14.

[287] Hogwood, B. W., and L. A. Gunn. 1984. *Policy Analysis for the Real World.* Oxford, UK: Oxford University Press. P. 227.

stakeholders, and in the alignment of their beliefs and values. Independence and credibility of evaluation become central for its legitimisation as a tool of governance. The next section addresses the formalisation of evaluation through the establishment of rules and institutions for its performance.

2.2.3 The Institutionalisation of Evaluation

The adoption of evaluation tools has made evident the need to develop formal actions to guide its utilisation within public organisations. This includes establishing the objectives of the evaluative activity, the scope of its influence, and the allocation of resources directed to this purpose. In other words, it concerns setting criteria or formal rules for its development. According to Mark et al. an evaluation policy is a set of 'high-level rules embedded in the legislation that are used to guide the practice of evaluation.'[288] Trochim understands this concept as 'any rule or principle that a group or organization uses to guide its decisions and actions when doing evaluation.'[289] The core element that differentiates these conceptualisations is the formality or explicitness of the rules that guide evaluation. While Mark et al. consider that these rules can take the form of regulation or a legislation bill, Trochim is more flexible by recognising that an evaluation policy can arise from informal guidelines and implicit principles. Nonetheless, he emphasises the importance of having written evaluation policies, because this …

> … can be thought of as a type of communication mechanism […] a signal to the entire organization and its stakeholders, communicating what evaluations should be done, what resources expended, who is responsible, how they should be accomplished, and so on.[290]

The establishment of formal rules for evaluation is strictly related to what Varone et al. define as the institutionalisation of evaluation:

> Understood as a 'systematization' of the expected, if not compulsory, recourse to evaluation, which can also be measured by its level of implementation within public administrations, political bodies and policy networks.[291]

---

[288] Mark, M. M., L. J. Cooksy, and W. M. K. Trochim. 2009. "Evaluation policy: An introduction and overview." *New Directions for Evaluation* 2009(123):3-11. Pp. 4-5.
[289] Trochim, W. M. K. Ibid."Evaluation policy and evaluation practice."13-32. P. 16.
[290] Ibid. P. 17.
[291] Varone et al. propose an index for measuring the degree of institutionalization of evaluation,

The underlying idea of an evaluation policy is the way in which rules and guidelines 'are expressed in legislation, regulations, and expenditures.'[292] Even though evaluations can be carried out based on informal guidelines, the relevance of an explicit evaluation policy is making a statement about different variables that in practice can represent major challenges, e.g. dissemination. This gives evaluation more leverage as a priority topic in the public agenda; it is embedded in a wider range of policies directed to improving and strengthening policymaking, e.g. transparency policies.

The institutionalisation of evaluation can have different meanings according to the elements considered important by the stakeholder in charge of its design. There is no consensus about what an evaluation policy must include. The American Evaluation Association Policy Task Force, for instance, has recognised seven key areas worth mentioning (Table 2.2).[293]

---

considering variables such as organisations directed to evaluation and the existence of an epistemic community of evaluation. Varone, F., Jacob, S., and De Winter, L. (2005). Polity, Politics and Policy Evaluation in Belgium. Ibid., 11(3), 253-273. P. 255.

[292] Datta, L. E. 2009. "Golden Is the Sand: Memory and Hope in Evaluation Policy and Wvaluation Practice." *New Directions for Evaluation* 123:33-50. P. 35.

[293] The World Bank, for example, considers that a successful monitoring and evaluation system should focus on three dimensions: utilization, quality and sustainability. Mackay, K. 2007. "How to build M & E systems to support better government." Pp. ix, 157 p. Washington, D.C., US: The World Bank. Pp. 23-24.

Table 2.2 Areas of an Evaluation Policy

Evaluation definition	How, if at all, is evaluation defined in an agency or in legislation? In such contexts, how is evaluation formally distinguished from or related to other functions such as program planning, monitoring, performance measurement or implementation?
Requirements of evaluation	When are evaluations required? What programs or entities are required to have evaluations? How often are evaluations scheduled? What procedures are used to determine when or whether evaluation takes place?
Evaluation methods	What approaches or methods of evaluation are recommended or required by legislation or regulation, for what types of programs or initiatives?
Human resources regarding evaluation	What requirements exist for people who conduct evaluations? What types of training, experience or background are required?
Evaluation budgets	What are the standards for budgeting for evaluation work?
Evaluation implementation	What types of evaluation implementation issues are guided by policies? For instance, when are internal versus external evaluations required and how are these defined?
Evaluation ethics	What are the policies for addressing ethical issues in evaluation?

Source: Based on American Evaluation Association, 2007. AEA Evaluation Policy Task Force Charge.

These areas underline the elements of the evaluation process that require guidance or control from a particular stakeholder; usually it is the government (central or local) who is in charge of this task. Core aspects of the institutionalisation of evaluation are vulnerable to political influence. As it will be discussed later, the extent to which stakeholders can exercise this control depends on the characteristics of the institutional framework.[294] The evolution of an evaluation policy can be a gradual process, transiting from a set of informal guidelines to an integrated and regulated policy instrument.[295] This was observed by Leeuw in the case of the Netherlands:

> In the 1980s, evaluation blossomed in the Netherlands, and informal evaluation policy not only was formulated but began to evolve into formal policy. As described below, this dramatic transition can be attributed to four
> factors: (1) demands by Parliament, (2) new vigor in the NAO, (3) the New

---

[294] See: Bamberger, M. 1991. "The Politics of Policy Evaluation in Developing Countries." *Evaluation and Program Planning* 14(4):325-39.

[295] See: Castro, M. F., G. Lopez-Acevedo, G. Beker Busjeet, and X. Fernandez-Ordonez. 2009. "Mexico's M&E System: Scaling Up from the Sectoral to the National Level." Washington, D.C., US: The World Bank.

Public Management movement, and (4) the Dutch administrative culture.[296]

The successful implementation of an evaluation policy is multifactorial, it includes political commitment and support, as in the case of the Netherlands, but also the development of evaluation capacities. In the case of the European Union, for instance, fostering the adoption of evaluation tools in member countries was also an incremental process in which evaluation was promoted …

> … through transition funding before and soon after EU accession, explicitly targeted at capacity development including training of evaluators; by sponsoring "twinning arrangements" between experts in evaluation (among other specialties) from established member states and their opposite numbers among the new members; by organizing evaluation network meetings in Brussels, nationally or at the regional level; and by providing guidance material and advice, including a series of guides to evaluation thinking and practice.[297]

Addressing the institutionalisation of evaluation in the context of this research matters for two main reasons. The first one is that the creation of rules and procedures for evaluation impacts the interaction among stakeholders and, as a result, the outputs (*how are these going to be generated?*) and outcomes *(how are these going to be used/disseminated?)* of this activity. The second is that the identification of core elements of the institutional framework provides a more effective lens for observing and analysing political influence. The hypothesis of this research suggests that the institutional framework can affect the extent to which politics can influence evaluation. Consequently, it is important to understand how the rules established for evaluation are transformed by political pressures.

The work of Chelimsky is useful for this purpose. She identifies three levels of political pressure that act on evaluation policies in the context of governmental structures (Figure 2.4).[298]

---

[296] Leeuw, F. L. 2009. "Evaluation policy in the Netherlands." *New Directions for Evaluation* 2009(123):87-102. P. 90.

[297] Stern, E. Ibid."Evaluation Policy in the European Union and its Institutions." 123:67-85. Pp. 69-70.

[298] Chelimsky, E. Ibid."Integrating Evaluation Units into the Political Environment of Government: The Role of Evaluation Policy."51-66.

Figure 2.4 Political Pressures to Evaluation in Government

Cross-branch politics	Bureaucratic climate	Dominant professional culture
This is the lofty arena in which the executive and legislative branches vie for power, engage in political partisanship, and direct ramified but wellaimed strikes on evaluative processes, outcomes, and reporting— especially reporting	This is a climate in which agency welfare, not public truth telling, is the primary political concern, and where evaluators may experience pressures that run the gamut from the forced embrace of bureaucratic norms and values to alteration of evaluation findings that could embarrass the agency	This culture tends to precipitate political clashes with the evaluation culture, sometimes because of unfamiliarity (when the unit is new), sometimes because of differences in theory or practice, but most often from a sense of rivalry or resentment

Source: Based on Chelimsky, E., 2009. "Integrating Evaluation Units into the Political Environment of Government: The Role of Evaluation Policy." New Directions for Evaluation, No. 123, pp 51-66.

Chelimsky concentrates on the politics of evaluation found within the governmental sphere. However, as analysed in Chapter 6, political pressures can come from many other different sources, i.e. groups of interests and political parties. These levels are consistent with the discussion about governance. The transformations of the state have affected not only the relationship between government and other stakeholders, but also within the governmental apparatus itself. Evaluation units are not only the means through which evaluation is institutionalised; they are also immersed in the dynamics of power and negotiation with evaluated bodies on the one hand, and with evaluators on the other. This role as mediator is a safeguard for preserving the independence and credibility of evaluation. Nonetheless, it cannot be ignored that these are not external or isolated bodies; they are, as any other agency, involved in the politics of the public sector. They can be perceived both as an aid for the process or as a menace for the interests of stakeholders.

Institutionalising evaluation is a way to give this activity more structure and coherence, but as Contandriopoulos states:

> It is not enough that the demand for efficiency be perceived by public opinion as sufficiently legitimate to have evaluation institutionalized; rather, the various actors must incorporate this new social standard into their ways of conceiving their own responsibilities and those of others.[299]

This shows that evaluation culture also represents a substantial element in the institutionalisation of evaluation. It is undeniable that promoting clear rules is necessary for incorporating evaluation in the logic of the public sector and for having more elements to cope with its political nature. The identification of key rules for the functioning of evaluation gives a more precise idea of those areas that are more susceptible to politics. The next section exposes the core discussions found in the evaluation studies literature about the hidden politics of evaluation.

2.3 The Hidden Politics of Evaluation

In the context of the study of the hidden politics of evaluation, it is necessary to deepen into the concept of politics, which has been studied extensively in the literature.[300] From Aristotle to contemporary thinkers, politics has represented an obsession for scholars from various disciplines, although no consensus about its definition has been reached. Lasswell, for example, understands politics as *who gets what, when, and how.*[301] This apparently simple definition has nurtured numerous debates around the multidimensionality of politics and its interpretation.[302] In his work, *Why We Hate Politics,* Hay identifies a dozen conceptualisations of politics. Among the most relevant are:

1. Politics as any and all social interaction occurring within the sphere of government;
2. Politics as government, where government is understood as a formal

---

[299] Contandriopoulos, A. P. 1999. "Is the institutionalization of evaluation sufficient to guarantee its practice?" *Cadernos de Saúde Pública* 15:253-56. P. 253.

[300] See Aristotle. 1995. *Politics.* Oxford, UK: Clarendon, Flinders, M. (Ed.). 2009. *The Oxford Handbook of British Politics.* Oxford, UK: Oxford University Press, McAnulla, S. 2006. *British Politics: A Critical Introduction.* London, UK: Continuum, Tansey, S. D. 2008. *Politics: The Basics.* London, UK: Routledge.

[301] Lasswell, H. D. 1970. "The emerging conception of the policy sciences." *Policy Sciences* 1(1):3-14.

[302] Farr, J., J. S. Hacker, and N. Kazee. 2008. "Revisiting Lasswell." Ibid.41:21-32, Marvick, D. 1980. "The Work of Harold D. Lasswell: His Approach, Concerns, and Influence." *Political Behavior* 2(3):219-29.

decision-making process the outcomes of which are binding upon member
of the community in question;
3. Politics as public and formal set of processes and rituals through which
the citizens of state may participate, often at arm's length, in the process of
government;
[...]
6. Politics as a process of public deliberation and scrutiny of matters of
collective concern or interest to a community;
7. Politics as a process for holding to account those charged with
responsibility for collective decision-making in the community.
[...].[303]

Looking at politics as a means for scrutinising those with the power to influence in the public sphere, e.g. government, helps explaining the role of evaluation in policymaking. This demands looking carefully at the relationships between those stakeholders involved and what this means for the scrutiny power vested to evaluation. Politics, therefore, is understood here as the way in which stakeholders use their resources (including power) to participate and influence the process and outcomes of evaluation according to their own agendas.

The transition experienced in the conceptualisation of evaluation across time is an important element to highlight. From the ideas about the need for objective and 'value-free' evidence obtained through robust methodologies, scholars have moved towards the impossibility—and undesirability—of abstracting evaluation from its political context.[304] Since the early seventies, academics and practitioners have begun debating about the interaction between evaluation and politics.[305] At that time, most of the discussion focused on the pressures faced by evaluators in the field and the reticence of operators to cooperate with the process. Since then, the discussion has reached other levels, but it is still at an early stage in the development of new knowledge, particularly in terms of empirical evidence. Taking this into account, the contribution of this thesis can be summarised in three major points:

---

[303] Hay, C. 2007. *Why we hate politics*. Cambridge, UK: Polity Press. Pp. 61-62.
[304] See Krathwohl, D. R. 1980. "The Myth of Value-Free Evaluation." *Educational Evaluation and Policy Analysis* 2(1):37-45, Scriven, M. 2007. "The logic of evaluation." Edgepress.
[305] See Palumbo, D. J. 1987. *The Politics of Program Evaluation*. California, US: SAGE, Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106, ibid., —. 1993. "Politics and Evaluation: A Reprise with Mellower Overtones." *American Journal of Evaluation* 14(1):107-09.

- The study of the political nature of evaluation has been constrained to the evaluation studies field. Consequently, little is known about the implications for the broader context of the policy process and governance. This thesis aims to produce a more comprehensive analysis of what the politics of evaluation does mean and how the current conditions in which this activity takes place affect its process and outcomes.
- The hidden politics of evaluation require a more integrated and systematic framework for grasping those key factors of evaluation that are affected. This thesis proposes a thematic framework to guide this discussion and serves as a basis for analysing this topic both theoretically and empirically.
- Most of the knowledge generated about the politics of evaluation has focused on explaining *how* this affects the practice of evaluation. However, *why* politics affects evaluation has been little explored. In this sense, the emphasis given by this thesis to the institutional framework aims to fill this knowledge gap by discussing those elements related to the rules and procedures for evaluation that affect the relationships that are established as a consequence of this activity.

An additional contribution is the comparative approach adopted by this research. Looking at the institutional architecture for evaluation in two different countries is useful for understanding how social, political, and administrative factors lead to a different set of rules for evaluation and, consequently, to a different way in which politics are manifested.[306]

2.3.1 General Notions About the Hidden Politics of Evaluation

The early development of the knowledge about the politics of evaluation was mostly based on the perspective of the evaluators, who identified the rationale underpinning the commissioning of evaluations and the values at stake.[307] They faced the conflict of preserving the values associated with their role as researchers (independence, credibility, etc.), while promoting evaluation as a more influential activity in policymaking.[308]

One of the most valued contributions is the work of Weiss, who identifies three key ways in which politics affects evaluation:

> First, the policies and programs with which evaluation deals are the creatures of political decisions [...] Second, because evaluation is

---

[306] The work of Weaver and Rockman is useful for understanding the implications of institutional differences. Weaver, K. R., and B. A. Rockman. 1993. *Do Institutions Matter? Government Capabilities in the United States and Abroad*. Washington, D.C., US: The Brookings Institution.

[307] Datta, L. E. 2011. "Politics and Evaluation: More Than Methodology." *American Journal of Evaluation* 32(2):273-94.

[308] Patton, M. Q. 1986. *Utilization-focused evaluation*. London, UK: SAGE.

undertaken in order to feed into decision-making, its reports enter the political arena [...] Third, and perhaps least recognized, evaluation itself has a political stance [...] it makes implicit political statements about such issues as the problematic nature of some programs and the unchallengeability of others, the legitimacy of program goals and program strategies, the utility of incremental reform, and even the appropriate role of the social scientist in policy and program formation.[309]

Undeniably, Weiss's contribution is remarkable. However, her conceptualisation remains too short, considering the transformations that the state has suffered. Indeed, public policy is the result of political decisions and evaluation is embedded in dynamics wherein values and beliefs can be in conflict. Also the nature of evaluation itself increases its political stance, as there are now more stakeholders involved. It is no longer a bilateral relationship between evaluators and evaluands; civil society and pressure groups are now also involved in the scrutiny of policy. Politicians and government officials see the potential of evaluation findings as a means for legitimisation and political discrediting. Universities, research centres, and private consultants have a vested interest in participating in the market of evaluation. NGOs and charities are currently important producers of evidence to support or discard the statements made by evaluation. Moreover, the media has also found in evaluation an important source of information to highlight policy issues and to promote the public debate. This reality is much more complex than the one described by Weiss in the seventies; nonetheless, she set very significant theoretical grounds for the study of the politics of evaluation.[310]

The political component of evaluation can be seen both as an advantage and as a limitation or, as Datta accurately describes it, as 'a bad master but a good friend.'[311] As Hedrick explains, evaluation can benefit from political disagreements by supporting the conduction of studies by a third trustable party, i.e. independent agencies. On the contrary, it can generate bias in the

---

[309] Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106. P. 94.

[310] The work of Palumbo is also recognised as a substantial contribution to the field, he reflected about the different areas in which politics can interfere and, more importantly, what this meant for the *praxis* of evaluation. Palumbo, D. J. 1987. *The Politics of Program Evaluation*. California, US: SAGE.

[311] Datta, L. E. 2011. "Politics and Evaluation: More Than Methodology." *American Journal of Evaluation* 32(2):273-94. P. 274.

scope, implementation, and dissemination of evaluations.[312] Bovens et al. incorporate the political dimension in their definition of evaluation by seeing it as …

> … an inherently normative act, a matter of political judgement […] it can at best be informed but never fully dominated by scholarly efforts to bring the logic of reason, calculation, and dispassionate truth seeking to the world of policy making.[313]

For them, recognising the political nature of evaluation allows policymakers and politicians to make a better use of evaluation as a tool for the accomplishment of their objectives, e.g. legitimising a programme. More recently, the discussion has focused on the advocacy role of evaluators[314] and the political nature of evaluation instruments.[315]

The multiple stakeholders involved in policymaking, i.e. civil society, media, and political parties, have promoted the idea that evaluation can become an instrument for pursuing their interests. As Taylor and Balloch state: 'Different stakeholder views and interests expressed from positions of more or less power and these relations of power are capable of being independently grasped.'[316] This has fostered the development of evaluation approaches that incorporate this new dynamic. Responsive evaluation, for example:

> … focuses on stakeholder issues and engages stakeholders in dialogues about the quality of their practice. The aim is to heighten the personal and mutual understanding of stakeholders as a vehicle for practice improvement.[317]

The need to engage stakeholders in evaluation is not exempt from politics, especially considering that they are implicated in asymmetrical power

---

[312] Hedrick, T. E. 1988. "The Interaction of Politics and Evaluation." Ibid.9(3):5-14.

[313] Bovens, M., P. 't Hart, and S. Kuipers. 2006. "The Politics of Policy Evaluation." in *The Oxford Handbook of Public Policy*, edited by M. Moran, M. Rein, and R. E. Goodin. Oxford, UK: Oxford University Press. P. 319.

[314] Mohan, R. 2014. "Evaluator Advocacy: It Is All in a Day's Work." *American Journal of Evaluation* 35(3):1-7.

[315] Chelimsky, E. 2012. "Valuing, Evaluation Methods, and the Politicization of the Evaluation Process." *New Directions for Evaluation* 2012(133):77-83.

[316] Taylor, D., and S. Balloch (Eds.). 2005. *The Politics of Policy Evaluation*. London, UK: The Policy Press. P. 2.

[317] Abma, T. A. 2006. "The Practice and Politics of Responsive Evaluation." *American Journal of Evaluation* 27(1):31-43. P. 31.

relationships.[318] Key discussions about the political nature of evaluation concern the modification of its role within the public sector. Karlsson explains this as a:

> … shift from an evaluation structure whose function has been to deliver material that can be used in the building of various welfare programs, to an evaluative function with the purpose of feeding the ongoing administrative control and decision-making process.[319]

The evidence that evaluation generates can serve different stakeholders to support or disqualify a political decision.[320] This can explain why participatory approaches in evaluation have become more popular by recognising the existence of stakeholders and their potential influence.[321] Therefore, the connection between evaluation and politics can be seen from different perspectives. According to Taylor and Balloch, there are two approaches for understanding this relationship.[322] The first one is *rationalistic* (*scientific realism*)[323] which focuses on the values of objectivity and neutrality and is based on the assumption that evaluation is performed with scientific rigour and with no political influence.[324] Its purpose is to preserve evaluation from political pressures by ignoring the possible influence of politics and by constructing policy optimums that can be measured through evaluations.[325]

The second approach is *social constructivism* or *argumentative evaluation*, which assumes that there cannot be a division between politics and evaluation.[326] The

---

[318] Ibid.

[319] Karlsson, O. 2003. "Evaluation Politics in Europe: Trends and Tendencies." *Studies in Educational Policy and Educational Philosophy*. P. 9.

[320] See Rogers-Dillon, R. H. 2004. *The Welfare Experiments: Politics and Policy Evaluation*. California, US: Stanford Law and Politics.

[321] Bamberger, M. 1991. "The Politics of Policy Evaluation in Developing Countries." *Evaluation and Program Planning* 14(4):325-39.

[322] Taylor, D., and S. Balloch (Eds.). 2005. *The Politics of Policy Evaluation*. London, UK: The Policy Press.

[323] For further reference about realistic evaluation see Pawson, R., and N. Tilley. 1997. *Realistic Evaluation*. London, UK: SAGE.

[324] This vision 'suggests a technocratic or rationalistic fix to the problem of change.' Saunders, M. 2006. "Do We Hear the Voices?: The 'Presence' of Evaluation Theory and Practice in Social Development." *Evaluation* 12(2):251-64. P. 258.

[325] Bovens, M., P. 't Hart, and S. Kuipers. 2006. "The Politics of Policy Evaluation." in *The Oxford Handbook of Public Policy*, edited by M. Moran, M. Rein, and R. E. Goodin. Oxford, UK: Oxford University Press.

[326] This approach conceives 'policy-making as an ongoing dialogue, in which both governmental and societal actors contest their views on policy issues by exchanging arguments.' Van der Knaap, P. 1995. "Policy Evaluation and Learning: Feedback, Enlightenment or Argumentation?"

central argument is that knowledge derived from evaluations is subjective and contextual. For social constructivists, evaluation generates information built on the values and beliefs of the participants and needs to be systematically interpreted. In terms of the hidden politics of evaluation the changes promoted by evaluation are a consequence of the exchange of perspectives and opinions between stakeholders, as Valovirta observes:

> The reason for the importance of argumentation in public policy is simple: as the number of actors involved in decision making increases, the need for negotiation, discussion and argumentation increases.[327]

This perspective recognises the inherent political nature of evaluation and the dynamic when it occurs. While the behaviour of stakeholders can be highly unpredictable, the existence of institutional arrangements for evaluation helps in reducing uncertainty by providing them a clearer idea of how the process works (or is *supposed* to work). The hypothesis established in this thesis suggests that the institutional framework influences the interaction between politics and evaluation and the way in which an evaluation policy is designed and implemented. In line with this, Widmer recognises that 'the purpose and the (intended) utilization should be the starting point in defining the most appropriate institutional design.'[328] The value given to institutions as elements that constrain and to an extent predict the behaviour of stakeholders gains more importance, specifically in terms of how these institutions can promote the influence of evaluation.[329] As Henry notes, evaluation has transited from being usually constrained to the executive branch to be part of the legislative structure and even an activity performed by NGOs, as now, 'evaluation findings have to pass scrutiny from a variety of individuals representing numerous institutions that have a plethora of interests.'[330] This makes clearer the connection between evaluation and auditing as mechanisms of internal and external control. Consequently, the course of an evaluation relies on the

Evaluation 1(2):189-216. P. 189.

[327] Valovirta, V. 2002. "Evaluation Utilization as Argumentation." Ibid.8(1):60-80. P. 65.

[328] Widmer, T., and P. Neuenschwander. 2004. "Embedding Evaluation in the Swiss Federal Administration: Purpose, Institutional Design and Utilization." Ibid.10(4):388-409. P. 406.

[329] Gaarder, M. M., and B. Briceño. 2010. "Institutionalisation of Government Evaluation: Balancing Trade-Offs." International Initiative for Impact Evaluation (3ie).

[330] Henry, G. T. 2001. "How Modern Democracies Are Shaping Evaluation and the Emerging Challenges for Evaluation." *American Journal of Evaluation* 22(3):419-29. P. 422.

configuration of the different agreements established among stakeholders. As any other stage of the policy cycle, evaluation is limited by power and its distribution denotes a key element for uncovering the hidden politics of evaluation.[331] Power can be understood as the capacity of stakeholders to exercise their resources, i.e. financial, data, etc. to pursue those interests that might be affected by evaluation.[332]

Hoefer recognises two perspectives for understanding power in policymaking. The first states: 'Those with power resources, particularly money, are the ones who decide what government policy becomes.'[333] As many decisions in evaluation depend on the availability of resources, its control can represent a disguised strategy for affecting the course of an evaluation. The second suggests that the 'use of power is to keep a proposal from being adopted or considered'.[334] Power is used here to prevent or to provoke a particular outcome. In both cases, evaluation can give stakeholders the opportunity of both legitimating and supporting a policy or building arguments against it. In this last case, a stakeholder's statement about the convenience or inconvenience of maintaining a policy can be supported by evaluation evidence. This is consistent with the discussion about the legitimisation function of evaluation, even in those cases in which evaluations are intentionally carried out for supporting a predetermined position.[335] Nonetheless, this support depends on the extent to which the whole evaluation system is considered legitimate. As House points out:

> Evaluation persuades rather than convinces, argues rather than demonstrates, is credible rather than certain, is variably accepted rather than compelling.[336]

---

[331] In this thesis, power is understood as a relationship in which 'A has power over B to the extent that he can get B to do something that B would not otherwise do.' This responds mostly to the role of evaluation as a mechanism to promote changes within public organizations. The information produced by evaluation tools can contribute to shape the behaviour of the stakeholders involved towards a desired outcome. See: Dahl, R. A. 1957. "The concept of power." *Behavioral Science* 2(3):201-15. Pp. 202-203.
[332] Orr, S. K. 2010. "Exploring Stakeholder Values and Interests in Evaluation." *American Journal of Evaluation* 31(4):557-69.
[333] Hoefer, R. 2012. *Policy Creation and Evaluation: understanding welfare reform in the United States.* United States: Oxford University Press. P. 44.
[334] Ibid. P. 44.
[335] Banner, D. K. 1974. "The politics of evaluation research." *Omega* 2(6):763-74.
[336] House, E. R. 1980. *Evaluating with Validity.* London, UK: SAGE. P. 73.

This discussion provides a general panorama of what the politics of evaluation mean. More specifically, the theoretical debate stresses three major arguments worth highlighting. The first one is that evaluation is per se embedded in a political context, its object of study, the values underpinning it, and the potential consequences of its findings are affected by political decisions. The second argument is that the increasing number of stakeholders involved in evaluation inevitably generates political disputes derived from the multiple interests that intervene and the asymmetrical power relationships existent. The third is that a gap exists in our knowledge, i.e. scholars have highlighted the politics involved in evaluation and the need for practitioners and evaluators to be aware of the possible implications in their practice, but there is very little referral to the way in which the institutional framework affects the dynamics between politics and evaluation and why different institutional arrangements favour or limit this relationship. This chapter has mapped the broad intellectual canvas on which this thesis exists. It has therefore deployed a fairly broad brush in analytical and theoretical terms. The next section offers a more fine-grained approach by drawing upon this wider literature in order to design a five-part thematic framework that will then be applied in the later empirical chapters. The aim of this framework is to develop the analytical tradition and leverage of evaluation studies as related to political dynamics. The remainder of this chapter examines each theme in detail.

2.3.2 Thematic Framework for Understanding the Hidden Politics of Evaluation

This thesis proposes a thematic framework that provides a more structured guide for understanding the hidden politics of evaluation. It represents an original and relevant contribution to the existent knowledge in at least two senses. On the one hand, it provides a clearer conceptualisation of the way in which politics and evaluation interact. On the other hand, the identification of key variables related to this interaction facilitates its study within an empirical context.

Despite the intangibility of politics, this framework sets out key themes for analysing its relationship with evaluation that can serve as a basis for looking at different cases through a structured lens and for setting conditions for comparability. Table 2.3 presents this thematic framework based on five core elements: *purpose, resources, process, outputs*, and *outcomes*; each is developed in this section.

Table 2.3 Thematic Framework for the Understanding of the Hidden Politics of Evaluation

Theme	Meaning	Key Questions	Analytical Variables	Key Literature References
Purpose	The aims and ambitions of the evaluation process (legal obligation, legitimacy of public decisions, identification of goals achieved, etc.).	Which is the main purpose of the evaluation? Which are the expected results? Which will be the effect on the policymaking process?	Role given to evaluation (support to the policy process, legitimisation)	Bjørnholt and Larsen, 2014; Palumbo, 1987
			Object of evaluation	Vartiainen, 2002; Owen and Rogers, 1999
			Scope	Mark et al., 2009; Howell and Yemane, 2006; Hedrick, 1998
			Timing	Bamberger et al., 2012; Balthasar, 2009; Mohan and Sullivan, 2006
			Institutionalization	Mark et al., 2009; Trochim, 2009; Datta, 2009; Furubo et al., 2002; Contandriopoulos, 1999
Resources	The capacity of the evaluating organisation in terms of staff, finance, knowledge and political capital.	Is there reliable and pertinent information to perform evaluations? Do public agencies have the financial resources required to perform evaluations? Do public officials have the required skills to conduct and manage evaluations?	Funding	Fletcher and Dyson, 2013; Bamberger et al., 2012; Chelimsky, 2008; De Lay and Manda, 2004
			Data	Bamberger et al., 2012; Guenther, 2008; De Lay and Manda, 2004; Chelimsky, 1995; McLemore and Neumann, 1987
			Human Resources	Desautels and Jacob, 2012; Lopez-Acevedo et al., 2012; Leeuw et al., 1999; House, 1997
			Methodology	Chelimsky, 2012; Green et al., 2007; Patton, 1987
			Time restrictions	Salisbury et al., 2010, Mohan and Sullivan, 2006

Table 2.3 *Continued*

Theme	Meaning	Key Questions	Analytical Variables	Key Literature References
Process	The way in which evaluation studies are undertaken and the 'tools of government' at the evaluator's disposal.	Does the evaluation process promote independency and objectivity? Are there reliable instruments to evaluate public programmes and policies? Do the professional selected to perform evaluations are chosen on the basis of their skills? Do the politics of evaluation influence the process?	Interaction between stakeholders	Jacob, 2011; Baur et al., 2010 Chelimsky, 2009; Barbier, 1999
			Role of commissioners	The LSE GV314 Group, 2014; Palfrey et al., 2012; Chelimsky, 2009
			Independence	Brown and Klerman, 2012; Lonsdale , 2008; Chelimsky, 2008
			Credibility	Copestake, 2014; Mohan, 2014; Sturges, 2014; Bozeman, 2004
			Openness to external participants (democratization of evaluation)	Picciotto, 2015; Hanberger, 2006; Green, 2001; Ryan and Johnson, 2000
Outputs	The final assessment produced by the evaluator and the dissemination process through which it may be debated, assessed and acted-upon.	Does evaluation produce reliable, relevant and pertinent information about public policy? Are the results derived from evaluation used? Which are the mechanisms for the dissemination of results? Can it be identified the influence of politics in the outputs of evaluation?	Use	Neuman et al., 2013; Balthasar, 2009; Ginsburg and Rhett, 2003; Feinstein, 2002; Weiss, 1998b
			Dissemination	Green et al., 2009; Lawrenz et al., 2007

Table 2.3   *Continued*

Theme	Meaning	Key Questions	Analytical Variables	Key Literature References
Outcomes	The broader long-term impact and role of evaluation within the politico-administrative system.	Are there mechanisms to identify the outcomes of evaluation?	Effects in the policymaking process	Bogenschneider and Corbett, 2010; Geva-May, 2004
		Is it possible to identify evidence about the utilization of evaluation results and its effect on the improvement of policymaking?	Development of evaluation capacities	Heider, 2011; Rist et al., 2011; McDonald et al., 2003
		Is it possible to identify the long-term effects of evaluation in the political arena?	Linkage to accountability	Bovens et al., 2014; Cunill-Grau and Ospina, 2012; Hoefer, 2000; Chelimsky, 1996

Purpose

The design and planning of an evaluation includes setting its objectives, defining the questions to be answered, and establishing the potential target audience.[337] The *purpose* of evaluation determines the reasons for conducting an evaluation. Its omission generates 'a risk that the evaluation will focus on the wrong concerns, draw the wrong conclusions and provide recommendations which will not be useful for the intended users of evaluation results.'[338] As a first approach to the development of an evaluation, the purpose makes a statement about the political values and beliefs that will guide this activity. This is a privilege limited only to one or a few stakeholders and is one of the reasons why it is a key political issue. The actor who controls this variable has the power to define important elements of the evaluation, such as the programmes to evaluate and the timing. It influences both the expected contribution within government, e.g. for practitioners, as well as for other stakeholders, e.g. the media. Evaluation is embedded in a political context in which values and beliefs affect the decisions made at this stage.[339]

While evaluation can be adopted as a mechanism for supporting policymaking, it can also have a legitimising function. This is not a minor issue in terms of politics. From an instrumental perspective, evaluation can provide insights about the functioning of a programme and the way in which processes are turned into outputs.[340] As a legitimisation mechanism, different actors use evaluations to support solutions in which they already believe. This can be seen more clearly in the discussion about the adoption of an internal or external model of evaluation and the rules that determine its nature.[341] While some systems favour self-evaluation for different reasons, including a more detailed

---

[337] For further reference about evaluation design it is advisable to look at the work of Howell and Yemane who conducted in the United States a study to assess the design of 12 federal programmes. Their study identified several challenges in terms of design, for example, the need for establishing more effective mechanisms for the dissemination of findings. Howell, E. M., and A. Yemane. 2006. "An Assessment of Evaluation Designs: Case Studies of 12 Large Federal Evaluations." *American Journal of Evaluation* 27(2):219-36.

[338] My M&E. 2004. "Purposes of Evaluation."

[339] Palumbo, D. J. 1987. *The Politics of Program Evaluation*. California, US: SAGE.

[340] Berriet-Solliec, M., P. Labarthe, and C. Laurent. 2014. "Goals of evaluation and types of evidence." *Evaluation* 20(2):195-213.

[341] Conley-Tyler, M. 2005. "A fundamental choice: Internal or external evaluation?" *Evaluation Journal of Australasia* 4(1 & 2):3-11.

knowledge of the object of evaluation, others value the independence and credibility that an external evaluation is supposed to offer.[342] Values are prioritised by stakeholders and this affects how rules for evaluation are designed.[343] The object of evaluation depends also on the definition of the purpose. Evaluation can study a policy, a programme, an organisation, or even an individual.[344] This range of objects reflects the level of analysis and specificity that evaluation can reach.[345] It also affects the level of cooperation obtained for evaluation. Failures found at policy level might be considered less threatening that those at individual level. This finds echo in Hood's ideas about 'the blame game' in terms of the responsibility for the success or failure of a policy possibly being diluted in the process.[346]

The selection of the object can determine *a priori* the kind of information that will be produced, as 'the primary use of research is thought by some to be motivated by a desire to support prior positions rather than to inform decisions in the first instance.'[347] The wideness or narrowness of the scope is linked to the level of analytical depth. As a political decision, it can be constrained by the interests of stakeholders who can aim for the evaluation to look only at those variables that might lead to positive findings and avoid vulnerable areas, to 'address only one perspective, perhaps intentionally.'[348] An important question to address regards timing—*when to evaluate?*—which can be related to political convenience. For example, evaluations prior to electoral times can be a mechanism for making explicit the goals achieved during a term by a specific

---

[342] ASK Inyathelo. "Advantages and Disadvantages of Internal and External Evaluations."

[343] Contandriopoulos, A. P. 1999. "Is the institutionalization of evaluation sufficient to guarantee its practice?" *Cadernos de Saúde Pública* 15:253-56.

[344] Owen, J. M., and P. J. Rogers. 1999. *Program Evaluation. Forms and approaches*. Melbourne, Australia: SAGE.

[345] Vartiainen, P. 2002. "On the Principles of Comparative Evaluation." *Evaluation* 8(3):359-71.

[346] Hood describes the idea of blame in public organisations as 'the act of attributing something considered to be bad or wrong to some person or entity.' For this author, this implies two major elements: 'a perceived and avoidable harm or loss' and 'some attribution of agency—that harm was avoidable because it was caused by acts of omission or commission by some identifiable individual or organization.' Hood, C. 2011. *The Blame Game: Spin, bureaucracy, and self-preservation in government*. New Jersey, US: Princeton University Press. P. 6.

[347] Bogenschneider, K., and T. J. Corbett. 2010. *Evidence-Based Policymaking: Insights from Policy-Minded Researchers and Research-Minded Policymakers*. New York, US: Routledge. P. 2.

[348] Hedrick, T. E. 1988. "The Interaction of Politics and Evaluation." *American Journal of Evaluation* 9(3):5-14.

actor, e.g. party in office.[349]

All these issues are relevant for discussing the institutional framework. Rules for evaluation are 'a signal to the entire organization and its stakeholders, communicating what evaluations should be done, what resources expended, who is responsible, how they should be accomplished.'[350] These will shape the behaviour of stakeholders, including those setting the purpose.[351]

Blomberg and Waldo state that a clear definition of the purpose and the stakeholders' commitment to it can help evaluation in coping with the political context.[352] If this commitment exists, stakeholders with different political positions are more likely to reach agreements on the basis that information is reliable. The predomination of a particular stakeholder in the evaluation process allows having control of those variables that might affect their interests by reducing the uncertainty associated with the possible outcomes of the evaluation.[353] This is why the confrontation between stakeholders 'requires political, ethical, and methodological negotiations that become intertwined with each other in the cauldron of program-focused decision-making.'[354] The conduction of these negotiations will be affected by the formal and informal rules existent, e.g. who can access this decision.

Resources

The planning stage involves deciding about the availability and allocation of *resources* for the performance of evaluations. Regarding this issue, Bamberger identifies some key political issues of evaluation in developing countries related to the control over the evaluation process and the way in which resource

[349] Balthasar, A. 2009. "Institutional Design and Utilization of Evaluation: A Contribution to a Theory of Evaluation Influence Based on Swiss Experience." *Evaluation Review* 33(3):226-56.

[350] Trochim, W. M. K. 2009. "Evaluation policy and evaluation practice." *New Directions for Evaluation* 2009(123):13-32. P. 17.

[351] Mark, M. M., L. J. Cooksy, and W. M. K. Trochim. Ibid."Evaluation policy: An introduction and overview."3-11.

[352] Blomberg, T. G., and G. P. Waldo. 2002. "Evaluation Research, Policy, and Politics." *Evaluation Review* 26(3):340-51.

[353] Hedrick, T. E. 1988. "The Interaction of Politics and Evaluation." *American Journal of Evaluation* 9(3):5-14.

[354] Scheirer, M. A. 1998. "Commentary: Evaluation Planning is The Heart of the Matter." Ibid.19:385-91. P. 385.

restrictions impact evaluation.[355]

Funding, for example, seems to be one of the most effective ways to influence the process. The sponsor has the right (explicit or not) to define the terms of the evaluation, the variables to consider, and the questions to answer. The power of sponsors affects the conditions for the evaluation and promotes the establishment of a sort of dominant relationship with evaluators. Seen as clients, sponsors can define the characteristics for the product. As House observes, in the case of drug evaluations, 'Academics are bound by detailed contracts with sponsors that limit independence in every way.'[356] Regardless of the sponsor of an evaluation, it is clear that its performance depends on the funding provided by this actor.[357]

International organisations, for example, 'must demonstrate the value of their investments to justify these expenditures and continue funding.'[358] If the benefit of evaluation activities is not clearly perceived, *Which will be the incentive for allocating resources to this end?* This concerns the symbolic value given to evaluation. Aside from its potential benefits for policymaking and for accountability, evaluation can also be seen as a politically profitable activity associated with an image of commitment to values like transparency and efficiency.[359] Consequently, even in a hypothetical scenario of not perceiving any instrumental benefit from evaluation, it can still be an issue in the government's agenda because of its political value.

The relationship between evaluation and funders presents other political consequences. According to O'Brien et al., 'Managers seek to present themselves

---

[355] Bamberger performed a comparative analysis about politics in policy evaluation between the United States of America and developing countries. As a result, he proposed a framework for the discussion of politics in policy evaluation in developing countries; however, the variables he selected can be applied in the case of any country. Bamberger, M. 1991. "The Politics of Policy Evaluation in Developing Countries." *Evaluation and Program Planning* 14(4):325-39.
[356] House, E. R. 2008. "Blowback: Consequences of Evaluation for Evaluation." *American Journal of Evaluation* 29(4):416-26. P. 419.
[357] Chaplowe, S. G., and R. B. Engo-Tjéga. 2007. "Civil Society Organizations and Evaluation: Lessons from Africa." *Evaluation* 13(2):257-74.
[358] De Lay, P., and V. Manda. 2004. "Politics of Monitoring and Evaluation: Lessons from the AIDS Epidemic." *New Directions for Evaluation* 2004(103):13-31. P. 19.
[359] Radaelli, C. M. 1995. "The role of knowledge in the policy process." *Journal of European Public Policy* 2(2):159-83.

and their organizations in the most expedient light in particular to funding bodies.'[360] Evaluation can therefore become a means to pursue a better allocation of resources based on evidence about a policy, e.g. efficiency or effectiveness. It can represent a bargaining chip between public organisations and funding agencies.[361] As Fletcher and Dyson accurately state: "We live in a political climate that requires 'evidence' in exchange for funding."[362] This linkage between positive evaluation findings of a programme and the allocation of its budget was also identified by Chelimsky.[363] The control of funding is a way to increase the power of stakeholders by giving them more favourable conditions to negotiate.[364] This resonates with the question about how evaluation evidence is produced and which conditions are put in place to generate relevant and useful information: *How do decisions about resources affect these conditions and, consequently, the outputs of evaluation systems?*

Data, for instance, can be an object of political manipulation at least in two senses. The first one concerns the access and disclosure of information needed by evaluators to make their judgements. The political context in which an evaluation takes place 'may impose constraints—sometimes going as far as the total restriction or classification of information—on evaluator's access to data.'[365] This can be an instrument of control for the protection of the stakeholders' interests. Evaluators can face important obstacles in obtaining the data, particularly when dealing with a politically sensitive policy.[366] According to Guenther, this resistance can be explained by two reasons: the fear that information will be used to the detriment of the evaluand and the organisational

---

[360] O'Brien, T., S. Payne, M. Nolan, and C. Ingleton. 2010. "Unpacking the Politics of Evaluation: A Dramaturgical Analysis." *Evaluation* 16(4):431-44. P. 432.

[361] Newcomer, K. E., and S. L. Caudle. 1991. "Evaluating Public Sector Information Systems: More Than Meets the Eye." *Public Administration Review* 51(5):377-84.

[362] Fletcher, G., and S. Dyson. 2013. "Evaluation as work in progress: Stories of shared learning and development." *Evaluation* 19(4):419-30. P. 421.

[363] Chelimsky, E. 2008. "A Clash of Cultures: Improving the "Fit" Between Evaluative Independence and the Political Requirements of a Democratic Society." *American Journal of Evaluation* 29(4):400-15.

[364] Azzam, T. 2010. "Evaluator Responsiveness to Stakeholders." Ibid.31(1):45-65.

[365] Chelimsky, E. 1995. "The Political Environment of Evaluation and What it Means for the Development of the Field: Evaluation for a New Century: A Global Perspective." Ibid.16(3):215-25. P. 221.

[366] McLemore, J. R., and J. E. Neumann. 1987. "The inherently political nature of program evaluators and evaluation research." *Evaluation and Program Planning* 10(1):83-93.

culture that straitens people from disclosing information with external parties.[367]

The identification of multiple sources of information to develop feasible and pertinent research questions is central for the analysis.[368] Particularly in complex areas like health policy, the collection and systematisation of information demands the participation of multiple agencies and governmental bodies.[369] Consequently, actors need to establish connections to access different sources of information. They are therefore obliged to negotiate throughout the process to set the conditions for the evaluation. This concerns not only having resources, but also defining how these are managed and how stakeholders would be able to establish connections that might facilitate their access to data.[370] The growth of these relationships depends on the trust between stakeholders and on their political connections.[371] Cooperation, as discussed in the section about governance, becomes a key element.

Another important issue is the potential political manipulation of data, e.g. the suppression or modification of information potentially harmful for stakeholders. De Lay and Manda observed this in their study about the HIV/AIDS crisis in Africa:

> Data may be inaccurate, attributed to the wrong population, or threatening to political leaders [...] Those engaged in M&E have a professional and ethical responsibility to clearly disclaim the limits of data sources and the analyses based on them to ensure that the scale and scope of problems are as accurately portrayed as possible..[372]

Data can become an important element of political control in situations when

---

[367]Guenther, J. 2008. "The Role of the Evaluator in Remote Contexts." in *Public Seminar*. Alice Springs: Centre for Remote Health.

[368] This is one of the issues discussed in the case of the Cancer Strategy (Chapter 6).

[369] Saunders, R. C., and C. A. Heflinger. 2004. "Integrating Data from Multiple Public Sources. Opportunities and Challenges for Evaluators." *Evaluation* 10(3):349-65.

[370] Williams, D. D., and M. L. Hawkes. 2003. "Issues and Practices Related to Mainstreaming Evaluation: Where Do We Flow From Here?" *New Directions for Evaluation* 2003(99):63-83.

[371] The work of Norris is particularly relevant for understanding the relationship between trust and evaluation as 'evaluation depends for its success on cooperative relationships and a measure of trust.' Norris, N. 2007. "Evaluation and Trust." Pp. 139-53 in *Dilemmas of Engagement: Evaluation and the New Public Management*, edited by S. Kushner and N. Norris: JAI Press Inc.

[372] De Lay, P., and V. Manda. 2004. "Politics of Monitoring and Evaluation: Lessons from the AIDS Epidemic." *New Directions for Evaluation* 2004(103):13-31.P. 25.

access to crucial information is delayed, obstructed, or even denied. This uncovers that the hidden politics of evaluation as controlling data is a way to cope with a potential threat to the power of stakeholders. However, considering the conditions that governance imposes, it could be counterproductive to adopt a reluctant position to evaluation. Public officials may resist sharing information with evaluators, fearing that this might affect them individually, i.e. with job security,[373] but at the same time data can become an instrument for increasing their influence in the process. As Kniker explains, recalling his experience as chief of evaluation at the U.S. Department of State, 'The more data were shared with these customers and stakeholders, the more they came to rely and depend upon it, gradually making us an invaluable resource.'[374] Both the use of data for/from evaluation can be subject to politics. On the one hand, controlling the access to key data for the evaluation can influence its outputs. On the other hand, it can give more leverage to stakeholders to support the decision-making process and to reach a more privileged position.

Human resources for evaluation studies are also susceptible to politics. The development of evaluation capabilities within public organisations is recognised as a key element for an effective evaluation system.[375] It influences the level of power and discretion that stakeholders have in the process.[376] This relates back to the institutional framework. Responsibilities for the conduct of an evaluation policy can determine the power that stakeholders will have within government, i.e. finance/planning units *vis à vis* specialised evaluation agencies.[377] Investing in evaluation capabilities is a political decision. More knowledge also means more consciousness about its potential use in the public

---

[373] This issue relates with the resistance to evaluation that has been largely addressed by the literature. For further reference see: Geva-May, I., and W. Thorngate. 2003. "Reducing Anxiety and Resistance in Policy and Programme Evaluations. A Socio-Psychological Analysis." *Evaluation* 9(2):205-27, Taut, S., and D. Brauns. Ibid."Resistance to Evaluation. A Psychological Perspective." (3):247-64.

[374] Kniker, T. 2011. "Evaluation Survivor: How to Outwit, Outplay, and Outlast as an Internal Government Evaluator." *New Directions for Evaluation* 2011(132):57-72. P. 68.

[375] Mackay, K. 2007. "How to build M & E systems to support better government." Pp. ix, 157 p. Washington, D.C., US: The World Bank.

[376] Raina, R. S. 1999. "Patronage and Evaluation in the Indian Council of Agricultural Research." *Evaluation* 5(3):278-302.

[377] Lopez-Acevedo, G., P. Krause, and K. Mackay (Eds.). 2012. *Building Better Policies. The Nuts and Bolts of Monitoring and Evaluation Systems.* Washington, D.C., US: The World Bank.

sector and the idea of the smarter state discussed in this thesis.[378] A more professionalised evaluation body might impact the level of engagement of stakeholders because it concerns the modification of values, beliefs, and routines in organisations.[379]

The other issue refers to the market of evaluators which has been developing at different rates during the last decades. In the early 2000s, Donaldson and Scriven stated that evaluation was 'in a second major boom period in its rather short history' produced by a wide range of policy areas that demanded the participation of evaluators.[380] In Europe, for example, 'The number of consultancy firms active in the field of evaluation has risen sharply in all countries'.[381] This relatively recent blooming of evaluation is a consequence of the governments' financial and legitimacy crises and of a stronger demand for accountability.[382] At all levels (supranational, national, and local), the rise of evaluation generated a market of knowledge. In the context of the evidence-based policymaking approach, as Bogenschneider and Corbett observe, there are *knowledge- producers* and *knowledge consumers* interacting on a regular basis. The information derived from evaluation represents an asset in this market.[383]

This market is composed by public agencies that commission evaluation studies in exchange of a payment. This does not seem any different from the purchase of any other product in any other market. However, the way in which actors interact does differ. The number of buyers and sellers is reduced and prices cannot be determined by the market due to this lack of competiveness. As House identified in the case of the United States:

---

[378] The barriers to evaluation capacity building have been addressed by Lennie, J., J. Tacchi, M. Wilmore, and B. Koirala. 2015. "A holistic, learning-centred approach to building evaluation capacity in development organizations." *Evaluation* 21(3):325-43.

[379] Milstein, B., T. J. Chapel, S. F. Wetterhall, and D. A. Cotton. 2002. "Building Capacity for Program Evaluation at the Centers for Disease Control and Prevention." *New Directions for Evaluation* 2002(93):27-46.

[380] Donaldson, S. I., and M. Scriven. 2003. *Evaluating Social Programs and Problems: Visions for the New Millennium*. London, UK: Erlbaum. P. ix.

[381] Leeuw, F. L., J. Toulemonde, and A. Brouwers. 1999. "Evaluation Activities in Europe: A Quick Scan of the Market in 1998." *Evaluation* 5(4):487-96. P. 488.

[382] Pollitt, C. 1998. "Evaluation in Europe: Boom or Bubble?" Ibid.4(2):214-24.

[383] Bogenschneider, K., and T. J. Corbett. 2010. *Evidence-Based Policymaking: Insights from Policy-Minded Researchers and Research-Minded Policymakers*. New York, US: Routledge.

> There are only a few buyers – the government agencies – [...] that contract
> out evaluations, so the market is imperfect on the demand (government)
> side, the market is also imperfect on the supply (contractor) side [...]
> typically, only a few firms bid for contracts in most instances.[384]

A reduced universe of public agencies commissioning evaluations acts according to the conditions set by the legal framework; for example, the public bidding of evaluation contracts. The situation in the supply side is similar. Evaluation requires specialised expertise and proficiency in the use of methodological tools; *ergo,* very few people can access the market. The increasing demand for the development of research has contributed to the growth of the 'business of evidence' sector.[385] Here, it is worth reflecting upon how this situation has modified the behaviour of evaluators. Academics have research interests to pursue, but as service providers, these interests might shift and produce bias.[386]

As an economic activity, evaluation implies a struggle between the interests of the buyer (public agencies) and the seller (evaluation firms). The dynamics of these relationships of power influences how evaluation is undertaken.[387] This concerns directly the institutional framework as it has set boundaries and control mechanisms for their interaction. For example: *To what extent are evaluators able to perform their job without interference from politicians and public officials?* Independence arises here as an important variable*,* especially in relation to the nature and strength of the relationships that are established.[388] The idea of looking at the political influence in the management of resources in an evaluation is underpinned by continuous negotiation between evaluators and sponsors in terms of independence.[389] This will be discussed when

---

[384] House, E. R. 1997. "Evaluation in the Government Marketplace." *American Journal of Evaluation* 18(1):37-48. P. 38.

[385] See: PWC. 2012. "The Business of Evidence. A report prepared for the Market Research Society." London, UK.

[386] Luo, H. 2010. "The Role for an Evaluator: A Fundamental Issue for Evaluation of Education and Social Programs " *International Education Studies* 3(2):42-50.

[387] Jacob analysed the role of key actors in evaluation; particularly in the case of the client (commissioner) he highlights the political power associated to this actor, in terms of its capacity to define the focus and scope of an evaluation, the selection of the evaluator, among other relevant decisions. Jacob, S. 2011. "Sharing Power among Evaluation Players: Mission Possible?" in *Evaluation. Seeking Truth or Power?*, edited by P. Eliadis, J. E. Furubo, and S. Jacob. New Jersey, US: Transaction Publishers.

[388] Lonsdale, J. 2008. "Balancing Independence and Responsiveness: A Practitioner Perspective on the Relationships Shaping Performance Audit." *Evaluation* 14(2):227-48.

[389] Chelimsky, E. 2008. "A Clash of Cultures: Improving the "Fit" Between Evaluative Independence and the Political Requirements of a Democratic Society." *American Journal of*

addressing the process of evaluation.

Methodology represents also a technical resource susceptible to politics. The tools selected for the collection, systematisation, and analysis of data for evaluation will affect both the process and the outputs. The institutional framework can also influence the methodological selection by permitting the prerogative of this decision to those stakeholders in a privileged position. In some cases, this decision is taken by the commissioner; in others, evaluators can freely select the tools that they consider most suitable. The application of standardised instruments based on a unique methodology can also be an element defined by the institutional framework.[390] As Norris recognises:

> The increasing tendency of governments to prespecify the characteristics of good evaluation by providing guidelines and standards for their evaluation contracts stems from an understandable desire for greater predictability and control over the content and process of evaluation.[391]

Chelimsky states that the methodological selection is constrained by political factors. For instance, some agencies maintain strong commitments to certain methods, e.g. randomised controlled trials (RCT), generating some resistance to test other approaches.[392] This can be explained by different reasons, such as the preference of stakeholders or organisational routines, but it can also be a decision taken on the basis of selecting those methods that can deliver more benefits for a particular stakeholder, or that are more aligned with their expectations.[393] Chelimsky also addresses the case of the *single narrative* which refers to ...

> ... a narrative that sees increasing numbers of government programs and policies embodying a single idea, or positing a simple, one-on-one cause-and-effect relationship, both of which are established, not by evidence, but rather by suppressing existing evidence that is inconvenient to the

---

Evaluation 29(4):400-15.

[390] This is the case of the Mexican evaluation policy (see Chapter 5).

[391] Norris, N. 2005. "The Politics of Evaluation and the Methodological Imagination." *American Journal of Evaluation* 26(4):584-86. P. 584.

[392] Chelimsky, E. 2012. "Valuing, Evaluation Methods, and the Politicization of the Evaluation Process." *New Directions for Evaluation* 2012(133):77-83.

[393] For example, evaluation of the conditional-cash transfer programme *Oportunidades* in Mexico was benefited from using a RCT in a moment when strong arguments were needed to support the sustainability of the programme during the political transition. Faulkner, W. N. 2014. "A critical analysis of a randomized controlled trial evaluation in Mexico: Norm, mistake or exemplar?" *Evaluation* 20(2):230-43.

particular idea or relationship being advanced.[394]

This is an important threat derived from the control exercised over the process and from the perception of the potential benefits of certain methods.[395] This resonates with the *methodological prejudice* discussed by Patton, which occurs when evaluators are keen to use those methods more familiar to their discipline. This can be a conscious or unconscious way to bias evaluation and it can provoke rejection from other stakeholders and affect credibility.[396] This is a key factor because:

> Methodological credibility has great importance not just in the planning and conduct but also in the survivability and use of an evaluation. It is thus crucial that evaluators retain control of their methods choices [...] especially because of political pressures, which are a constant in government and against which they must eventually defend those choices.[397]

This discussion relates back to the concepts of *ideology and polarization* (Section 2.2.2). The beliefs and values of the stakeholders will affect the decisions made, and vice versa.[398] The level to which this can affect the evaluation process will depend on their power and resources, and on the conditions set by the institutional framework. A final resource to discuss is *time*. Constraints of this nature can interfere with the work of an evaluator and his/her capacity to perform a robust study. As Mohan and Sullivan suggest, this involves a negotiation between stakeholders for establishing a suitable time framework.[399] The control of time, usually by commissioners, is a political way to limit the scope of a study and to exercise power. Contracts for evaluation serve as a mechanism to define *ex ante* the conditions that will rule the relationship between commissioners and evaluators, including the time available for the analysis. Evaluators are constrained to work under this

---

[394] Chelimsky, E. 2012. "Valuing, Evaluation Methods, and the Politicization of the Evaluation Process." *New Directions for Evaluation* 2012(133):77-83. P. 77.

[395] Ibid.

[396] Patton, M. Q. 1987. "Evaluation's Political Inherency: Practical Implications for Design and Use." in *The Politics of Program Evaluation*, edited by Dennis J. Palumbo. California, US: SAGE.

[397] Chelimsky, E. 2007. "Factors Influencing the Choice of Methods in Federal Evaluation Practice." *New Directions for Evaluation* 2007(113):13-33. P. 30.

[398] Schwandt, T., and P. Dahler-Larsen. 2006. "When Evaluation Meets the 'Rough Ground' in Communities." *Evaluation* 12(4):496-505.

[399] Mohan, R., and K. Sullivan. 2006. "Managing the Politics of Evaluation to Achieve Impact." *New Directions for Evaluation* 2006(112):7-23.

framework without the possibility of expanding the focus of their work, affecting their methodological choices.[400]

According to Salisbury et al., time is a variable that creates tension because policy-makers need useful information in a relatively short time span, and evaluators can see this as 'risking running counter to methodological robustness.'[401] This will be discussed later in the outputs of evaluation. However, stakeholders will demand differentiated products according to their needs and interests. Politicians might give a higher value to practical findings that allow them to get advantage of a political momentum than to a robust academic paper with little practical application. This might produce conflicts among stakeholders,[402] but its solution is a matter of defining priorities, negotiating, and eventually, the predomination of someone's interests.

Process

Rules for the commissioning, conduction, and supervision of an evaluation provide stakeholders the power to influence this process. It is possible to identify two major institutional designs for evaluation (see Section 2.2.3). The first one gives the prerogative to public organisations, i.e. ministries, for carrying out internally or commissioning to an external party the development of evaluation. The second implies the existence of an evaluation agency that acts as an intermediary between evaluators and evaluands.[403] In both models, the relationships established can become extremely political because they are based on the resources and power of stakeholders, determining the extent to which they are willing to cooperate.

One of the most political connections occurs between commissioners and evaluators. This interaction leads to what House labels as 'socialization'

---

[400] Sturges, K. M. 2014. "External Evaluation as Contract Work: The Production of Evaluator Identity." *American Journal of Evaluation* 35(3):346-63. The LSE GV314 Group. 2014. "Evaluation under contract: Government pressure and the production of policy research." *Public Administration* 92(1):224-39.
[401] Salisbury, C., K. Stewart, A. Cameron, S. Peckham, M. Calnan, R. Lart, S. Purdy, and H. Thorp. 2010. "Making the Most of Policy Evaluations: Overview and synthesis of evaluations of the White Paper 'Our health, our care, our say'." Bristol, UK: University of Bristol. P. 110.
[402] Ibid.
[403] Balthasar, A. 2009. "Institutional Design and Utilization of Evaluation: A Contribution to a Theory of Evaluation Influence Based on Swiss Experience." *Evaluation Review* 33(3):226-56.

between public agencies and evaluation firms.[404] According to this author, both parties have privileged information that they might not be able to obtain in a more competitive market and that creates distortions that might influence the results. Major consequences of this 'socialization' can be differentiated in relation to the actor. Public agencies, for example, can benefit from finding out which firms are reliable and are able to manage their contracts. They can also identify who can successfully perform the work according to their standards and who is more likely to establish 'informal agreements' to fulfil the requirements set by the legal framework. As firms depend largely on the contracts provided by public agencies, they are willing to adapt their work team to fulfil their needs. What is more, this 'closeness' allows firms to know the administrative procedures of the public sector and to obtain relevant information about budget, relationship with other agencies, etc. This can lead to a bigger influence in their thoughts and beliefs, and bias in the results of evaluation.[405] Overall, excessive closeness entails the risk of affecting the objectivity of the process.

This can also be perceived in the pressure that public agencies exercise on evaluators. According to the LSE GV314 Group, in the UK this occurs at three major stages of the process: the design and planning stage, the management of the evaluation, and the report of the findings. Their study shows:

> The most effective constraint appears to be found when government specifies the nature of the research to be done at the outset. No other form of constraint has as powerful an effect on the degree to which the overall conclusions the researchers reach support government policy.[406]

What arises from this discussion is that as multiple interests are vested in evaluation, stakeholders will attempt to influence the process through different means. However, this cannot be too explicit as it would damage the credibility of a study and possibly the reputation of the stakeholder. Instead, influence is disguised as minor interventions, e.g. methodological decisions, data control,

---

[404] House, E. R. 1997. "Evaluation in the Government Marketplace." *American Journal of Evaluation* 18(1):37-48.
[405] House, E. R. (1997) *Op cit.*
[406] The LSE GV314 Group. 2014. "Evaluation under contract: Government pressure and the production of policy research." *Public Administration* 92(1):224-39. P. 236.

etc., that in the aggregate can affect the outputs.

Another interesting element of this relationship is the assumption of the costs of evaluation. This means: *Who pays for the development of a study?* Balthasar recognises it as an important factor that impacts the distance between evaluators and evaluands. As the institutional framework determines who will assume the cost of evaluations, this affects the content of the contract.[407] This exposes the relevance of the institutional framework in relation to the rules for the conduction of sub-processes of evaluation and the implications for independence through the control of payments. As discussed, independence and credibility are highly interiorised values, both in evaluation and auditing.[408] Lonsdale defines independence in evaluation as 'an ability to determine what, how and when work is undertaken without undue interference.'[409] It relies fundamentally on the capacity of evaluators to make a judgement based on objective evidence and to report it free from constraints or impositions of others. Nonetheless, this is not only a safeguard for the robustness of a study, but also it has political implications. Evaluation will produce political disagreements, for example, between executive agencies and legislative bodies, which need to be acknowledged.[410]

The discussion about internal or external evaluation[411] has been underpinned by the fallible idea that external evaluators will be more independent, ignoring the fact that the contracting-out process is not exempt from political influence.[412] As Gaarder and Briceño suggest:

---

[407] Balthasar, A. 2009. "Institutional Design and Utilization of Evaluation: A Contribution to a Theory of Evaluation Influence Based on Swiss Experience." *Evaluation Review* 33(3):226-56.
[408] See: International Federation of Accountants. 2010. "Handbook of International Quality Control, Auditing, Review, other Assurance, and Related Services Pronouncements." New York, USA.
[409] Lonsdale, J. 2008. "Balancing Independence and Responsiveness: A Practitioner Perspective on the Relationships Shaping Performance Audit." *Evaluation* 14(2):227-48. P. 228.
[410] Chelimsky, E. 2008. "A Clash of Cultures: Improving the "Fit" Between Evaluative Independence and the Political Requirements of a Democratic Society." *American Journal of Evaluation* 29(4):400-15.
[411] Pérouse de Montclos studies the different advantages and disadvantages of evaluations according to their degree of independence. See: Pérouse de Montclos, M. A. 2012. "Humanitarian action in developing countries: Who evaluates who?" *Evaluation and Program Planning* 35(1):154-60.
[412] Brown, A. B., and J. A. Klerman. 2012. "Independent Evaluation: Insights from Public

> The reality of their independence is determined mainly by who is funding them, and may be compromised if it is by the very managers in charge of the activities under evaluation. No undue influence needs to be exerted by the managers for the situation to be compromised, as consultants may be self-censoring to maintain their clients.[413]

The last point of these authors is pertinent to discuss the level to which the institutional framework can shape the behaviour of stakeholders. In fact, rules about whether evaluation is performed in-house or contracted out can be driven by the logic of creating more favourable conditions for independence, but it can also produce incentives for providing a more or less independent judgement according to the interests of stakeholders. There is a difference, as Weaver observes, between independence *de jure* and *de facto*. Even if the institutional framework establishes the creation of an independent evaluation unit, its independence can be compromised through the control of other variables, such as 'budget, staffing, and reporting structures'.[414] Here, evaluators have a key role because they rely on the preservation of their credibility.[415] However, it should not be forgotten that they might be interested in maintaining a good relationship with commissioners to assure future contracts, as 'evaluation is a business.'[416] This echoes Picciotto's reflection about evaluators acting more as consultants, as they 'favor a strictly advisory supporting role that does not challenge the implicit values of their employers or the process that underlie organizational performance.'[417]

Regarding credibility, Copestake understands this as …

> … one party being able to offer a sufficient combination of evidence and explanation to convince another party that a proposition is reasonable in the sense of being sufficiently plausible to act upon – not rational in a logical sense, perhaps, but neither irrational.[418]

Accounting." *Evaluation Review* 36(3):186-219.

[413] Gaarder, M. M., and B. Briceño. 2010. "Institutionalisation of Government Evaluation: Balancing Trade-Offs." International Initiative for Impact Evaluation (3ie). P. 6.

[414] Weaver, C. 2010. "The politics of performance evaluation: Independent evaluation at the International Monetary Fund." *Review of International Organizations* 5:365-85. P. 375.

[415] The LSE GV314 Group. 2014. "Evaluation under contract: Government pressure and the production of policy research." *Public Administration* 92(1):224-39.

[416] Brown, A. B., and J. A. Klerman. 2012. "Independent Evaluation: Insights from Public Accounting." *Evaluation Review* 36(3):186-219. P. 191.

[417] Picciotto, R. 2013. "Evaluation Independence in Organizations." *Journal of MultiDisciplinary Evaluation* 9(20):18-32. P. 19.

[418] Copestake, J. 2014. "Credible impact evaluation in complex contexts: Confirmatory and exploratory approaches." *Evaluation* 20(4):412-27. Pp. 416-417.

Stakeholders relate both independence and credibility to the legitimacy of this activity and the recognition of the trustworthiness of the findings. This can help improving not only policymaking but also its political sustainability.[419] Preserving these values gives stakeholders elements to cooperate according to their role. The linkage between evaluation and democratic values like transparency, inclusion, and accountability is considered a tool for including the views of the stakeholders involved in policy.[420] The democratisation of evaluation[421] has gained notoriety within governance, as 'Evaluation is now being judged not only on its outputs (reports findings, conclusions and recommendations), but also on its process in terms of actors' representativeness, inclusion and participation.'[422] The identification of the interests of stakeholders has become a central element and, as a result, negotiation is crucial for determining the level of power exercised by each stakeholder.[423]

According to House and Howe, the involvement of different stakeholders comprises inclusion, dialogue, and deliberation (elements linked to negotiation).[424] This approach is more based on the establishment of informal rules of interaction than on strict norms defined by the institutional framework. An example of this is the role that this perspective gives evaluators as facilitators or mediators in the negotiation of interests.[425] Associating evaluation with democracy resonates with its role as a risk-reduction

[419] Lustig, N. 2011. "Scholars who became practitioners. The influence of research on the design, evaluation and political suvival of Mexico's Antipoverty Program Progresa/Oportunidades." edited by Center for Global Development.

[420] Mathison, S. 2000. "Deliberation, Evaluation, and Democracy." *New Directions for Evaluation* (85):85-89.

[421] See: Greene, J. C. 1997. "Evaluation as Advocacy." *American Journal of Evaluation* 18(1):25-35, —. 2001. "Dialogue in Evaluation: A Relational Perspective." *Evaluation* 7(2):181-87, Greene, J. C., L. DeStefano, H. Burgon, and J. Hall. 2006. "An educative, values-engaged approach to evaluating STEM educational programs." *New Directions for Evaluation* 2006(109):53-71.

[422] Jacob, S. 2011. "Sharing Power among Evaluation Players: Mission Possible?" in *Evaluation. Seeking Truth or Power?*, edited by P. Eliadis, J. E. Furubo, and S. Jacob. New Jersey, US: Transaction Publishers. P. 67.

[423] See: Baur, V. E., A. H. G. Van Elteren, C. J. Nierse, and T. A. Abma. 2010. "Dealing with Distrust and Power Dynamics: Asymmetric Relations among Stakeholders in Responsive Evaluation." *Evaluation* 16(3):233-48. Azzam, T., and B. Levine. 2014. "Negotiating Truth, Beauty, and Justice: A Politically Responsive Approach." *New Directions for Evaluation* 2014(142):57-70.

[424] House, E. R., and K. R. Howe. 1999. *Values in Evaluation and Social Research*. London, UK: Sage.

[425] Barbier, J. 1999. "Inter-governmental Evaluation: Balancing Stakeholders' Expectations with Enlightenment Objectives?" *Evaluation* 5(4):373-86.

mechanism and as a tool for including different views in policymaking. As there is competition for resources and the behaviour of stakeholders is unpredictable, the formalisation of evaluation processes can help reducing uncertainty. In governance, as Picciotto suggests:

> Evaluation was conceived as a way to help identify and where necessary fill the democratic deficit embedded in public policies and programmes: limited transparency; weak accountability, limited participation by disadvantaged groups, disregard of social equity, etc.[426]

The contribution of evaluation to democracy as a mechanism that favours transparency and accountability is subject to the use given to findings and the acknowledgement of the circumstances in which it takes place.[427] In line with this, Stame suggests that in the context of the NPM reforms and governance, evaluation can provide evidence about the relationship between the changes that took place in policy and the variables that led to those changes.[428]

The democratisation of evaluation relies on the openness of the process to the different positions of stakeholders in terms of power and resources. On the one hand, the institutional framework can establish formal mechanisms for their participation. On the other hand, through informal agreements stakeholders can establish conditions for reaching a common purpose, taking advantage of the discretionary spaces found in the institutional framework. In any case, the key element is questioning *to what extent stakeholders are able to participate.* It is not the same degree of involvement to participate as an evaluator than as a member of a focus group. The level of participation allowed can be related to the level of power that stakeholders are willing to grant others and the mechanisms that the institutional framework has established for this purpose, i.e. barriers of access.

Outputs

The use given to the products of evaluation, i.e. reports, surveys, or studies, depends on the different arrangements that take place among stakeholders, as 'evaluation cannot be regarded as neutral; it holds power and seeks to fulfil its

---

[426] Picciotto, R. 2015. "Democratic evaluation for the 21st century." Ibid.21(2):150-66. P. 156.
[427] Hanberger, A. 2006. "Evaluation of and for Democracy." Ibid.12(1):17-37.
[428] Stame, N. Ibid."Governance, Democracy and Evaluation."7-16.

own interests.'[429] This thesis addresses two major issues: utilisation and dissemination. Considering asymmetries of power and resources, stakeholders can use evaluation outputs in different ways. The actor that controls the process is in a privileged position to promote, disseminate, or even dismiss the results of an evaluation according to his criteria and interests.[430] This advantage can be observed in the use of evaluation findings to support a pre-established political position. According to a study developed in Great Britain:

> The use of evidence tended to support a certain way of thinking about the world; a thought style which is produced through the unequal distribution of power and then plays a part in its reproduction.[431]

The legitimisation function of evaluation is used to validate a particular pre-conceived position. In some cases, since its inception the use that will be made of the findings has already been determined and it is linked to the user's preferences.[432] Neuman et al. emphasise the relevance of making a commitment to evaluation to promote utilisation,[433] seen as its ultimate goal. This commitment[434] is a central element of the participatory or democratic approach discussed before.[435] Different reactions from stakeholders to evaluation outputs can be observed, ranging from welcoming a report to an adverse reaction to the findings or even completely burying them. The negative reaction to evaluation can be explained by the anxiety generated by this activity,[436] while a positive one is linked to the identification of a particular benefit from the use of outputs.

---

[429] Furubo, J., and O. Karlsson. 2011. "Evaluation: For Public Good or Professional Power?" Pp. 240 in *Evaluation: Seeking Truth or Power?*, edited by P. Eliadis, J. E. Furubo, and S. Jacob. New Jersey, US: Transaction Publishers. P. 33.

[430] Guenther, J., E. Williams, and A. Arnott. 2010. "The politics of evaluation: evidence-based policy or policy-based evidence?" in *NARU Public Seminar Series*. Darwin, Australia.

[431] Stevens, A. 2011. "Telling Policy Stories: An Ethnographic Study of the Use of Evidence in Policy-making in the UK." *Journal of Social Policy* 40(02):237-55. P. 251.

[432] Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106.

[433] Neuman, A., N. Shahor, I. Shina, A. Sarid, and Z. Saar. 2013. "Evaluation utilization research—Developing a theory and putting it to use." *Evaluation and Program Planning* 36(1):64-70.

[434] See Lloyd, N., and L. Harrington. 2012. "The challenges to effective outcome evaluation of a national, multi-agency initiative: The experience of Sure Start." *Evaluation* 18(1):93-109, Mensa-Bonsu, I. F., and M. Holm Andersen. 2010. "Involving Developing Countries in the Evaluation of Development Cooperation: The Case of a Joint Evaluation of Development Cooperation between Ghana and Denmark." Ibid.16(3):295-308.

[435] The work of Greene focuses in the development of an effective dialogue in evaluation to incorporate the views, interests and values of stakeholders. Greene, J. C. 2001. "Dialogue in Evaluation: A Relational Perspective." Ibid.7(2):181-87.

[436] See Geva-May, I., and W. Thorngate. 2003. "Reducing Anxiety and Resistance in Policy and Programme Evaluations. A Socio-Psychological Analysis." Ibid.9:205-27.

As Weiss identifies, evaluation can be used as a means for legitimisation or even as 'an excuse for inaction.'[437] The political capital built around policies can influence the different levels of impact that evaluation will have and the leverage given to evaluation findings depends greatly on the ability of evaluators to communicate its relevance for decision-makers.[438] Therefore, the findings of the evaluation of a high-profile policy would tend to get more attention than those from more modest programmes.[439]

The second element regarding outputs is dissemination, which can also represent a political decision. This refers to 'the appropriateness of the means used to facilitate access to the evaluation'.[440] Stakeholders can facilitate or hinder the access to the audiences of evaluation, even presently, when technology can considerably simplify this process.[441] Commissioners or sponsors have an important role in dissemination because they usually have more influence at this stage. This does not necessarily refer to the deliberate manipulation of findings, but to the intentional suppression of reports that might be unfavourable for stakeholders' interests:

> If the findings are considered to be politically sensitive, they may not be publicly released or they may be ignored [...] a further possibility is that the department which commissioned the evaluation will restructure and the evaluation findings no longer have relevance.[442]

Weiss argues: 'If the findings are negative, officials may not completely bury the report (although sometimes they try), but they can at least release it with their own interpretations.'[443] Thinking of evaluation as a tool for risk-reduction, negative findings can be seen as a potential threat to their interests. Therefore: *What will be the incentive for stakeholders to disclose even that information is*

---

[437] Weiss, C. H. 1998b. "Have We Learned Anything New About the Use of Evaluation?" *American Journal of Evaluation* 19(1):21-33. P. 25.

[438] Blomberg, T. G., and G. P. Waldo. 2002. "Evaluation Research, Policy, and Politics." *Evaluation Review* 26(3):340-51.

[439] High-profile cases were the evaluations of *Sure Start* in the UK and *Oportunidades* in Mexico.

[440] Feinstein, O. N. 2002. "Use of Evaluations and the Evaluation of their Use." *Evaluation* 8(4):433-39. P. 434.

[441] Bozeman, B. 2004. "The Internet's Impact on Policy Evaluation: Information Compression and Credibility." *Evaluation Review* 28(2):156-74.

[442] Guenther, J., E. Williams, and A. Arnott. 2010. "The politics of evaluation: evidence-based policy or policy-based evidence?" in *NARU Public Seminar Series*. Darwin, Australia.

[443] Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106. P. 101.

considered potentially harmful? The LSE GV134 Group observed that evaluators in the UK constantly receive petitions for changing the wording of findings.[444] Even if rephrasing findings cannot be seen as an open attempt at manipulation, it does represent a source of pressure. Seeing it as a commercial rapport, it is the client asking for modifications; the client who is paying for the services and a potential source for future business. Consequently, there might be incentives for evaluators to be complacent.

The nuances of how politics affects the dissemination of evaluation outputs concern the level of control and the power of stakeholders. The inclusion of new participants in the process of evaluation is expected to make it more democratic. Chelimsky highlights the advocacy role of evaluators and the relevance of preserving their integrity:

> Although the policy use of findings will surely be reduced in such situations if opponents are powerful enough, still – once evaluations with important findings have been published, no matter how reluctantly – it is evaluators' duty to bring their knowledge forward.[445]

It is certain that a struggle of interests might take place at this stage. However, evaluators are not the only persons responsible for protecting findings from political interference. The establishment of institutional safeguards for this purpose and the promotion of evaluation culture could foster appropriate dissemination.

Outcomes

Outcomes, as the long-term effects of evaluation on the policymaking process, can be very difficult to observe because they entail making some assumptions that cannot always be supported empirically. The first one is that all evaluations are robust/high quality and that they produce relevant and pertinent findings. In reality, the establishment of criteria to measure the quality of evaluation can be subjective and inconclusive.[446] The second refers to how policymakers use

---

[444] The LSE GV314 Group. 2014. "Evaluation under contract: Government pressure and the production of policy research." *Public Administration* 92(1):224-39.
[445] Chelimsky, E. 1995. "The Political Environment of Evaluation and What it Means for the Development of the Field: Evaluation for a New Century: A Global Perspective." *American Journal of Evaluation* 16(3):215-25.P. 220.
[446] Forss, K., and J. Carlsson. 1997. "The Quest for Quality—Or Can Evaluation Findings Be

findings to support their tasks. The difficulties in isolating the net effect of evaluation impede measuring accurately its contribution.[447] As Rauschmayer et al. observe, establishing a causal link between inputs and outputs can lead to the misinterpretation of the contributions of evaluation:

> If we look at the consequences only, we fail to assess the degree to which a specific intervention is responsible for the changed situation. If we instead examine the direct outputs, i.e. the laws, rules, programmes, activities and budget allocations produced by the governance process, we are still unable to determine whether the outputs have had the intended effects.[448]

Boardman agrees by underlining the need for an integrated approach that looks at both the process and the outcomes of public policy.[449] This resonates with the discussion about the scope and purpose of evaluation: *Who will determine the areas that will be prioritised for evaluation? To what extent is this prioritisation driven by political reasons?* This will vary according to each stakeholder's interests. For politicians, for example, evidence about the outcomes of a policy will be more useful than that of the process. The focus on outcomes can represent an important element in the context of governance. In line with this, Rotberg argues that performance measurements allow distinguishing those elements in the government's machine that are working from those that are not.[450] This gives *de facto* more relevance to certain aspects of a policy than to others. Performance information, as an input for evaluation, can provide a clearer identification of criteria for measuring the performance of governments and can make it easier for evaluators to understand the logic of a programme. The focus on the value generated by public policy seems to be a consequence of the new approaches in which beneficiaries/clients of public services are at the centre of the process.[451]

Trusted?" *Evaluation* 3(4):481-501.

[447] The complexity of measuring public goods has been recognised by Fukuyama in the broader context of governance. For him, 'outputs like health or education are not simply the consequences of public action; the public sector interacts with the environment around it [...] to produce results.' Fukuyama, F. 2013. "What Is Governance?" *Governance* 26(3):347-68. P. 355.

[448] Rauschmayer, F., A. Berghöfer, I. Omann, and D. Zikos. 2009. "Examining processes or/and outcomes? Evaluation concepts in European governance of natural resources." *Environmental Policy and Governance* 19(3):159-73. P. 164.

[449] Boardman, C. 2014. "Assessing Governance: The Importance of Evaluating Policy Outcomes in National Mission Areas." *Governance* 27(3):519-26.

[450] Rotberg, R. I. Ibid."Good Governance Means Performance and Results."511-18.

[451] Schalock, R. L., and G. S. Bonham. 2003. "Measuring outcomes and managing for results." *Evaluation and Program Planning* 26(3):229-35.

The key element here is not to what extent evaluation influences policymaking, but the extent to which it has been able to permeate the logic of policymakers. Evaluation needs to find more effective ways into the political dynamics and to transcend its managerial role to become an element for the political debate. As Richardson states:

> [B]arriers to full and effective use of social science in public policymaking
> are related to a separation of knowledge production from its consumption
> and utilization, and by a lack of mechanisms to allow for the political nature
> of knowledge transfer and decision-making processes.[452]

The construction of arguments based on evidence from evaluation can generate more credibility due to the 'rational' and 'objective' values underpinning it.[453] Recognising the political context can give evaluation more leverage as 'public policy must be analyzed as a political outcome and as a force that influences political actors, organizes political understandings, and structures political relations.'[454]

This relates back to the development of evaluation capacities.[455] Commitment to evaluation can be seen in how governments devote resources and efforts to build evaluation systems. This means prioritising different values, i.e. independence, credibility, and utility.[456] These can fall into conflict with the agenda of the stakeholder conducting the evaluation system; yet, the development of evaluation capacities will provide stakeholders with more mechanisms to engage in evaluation under more favourable conditions. Fostering these capacities can depend on political will because it requires considering 'the effect of informal practices, symbolic actions, evaluative beliefs,

[452] Richardson, L. 2014. "Putting the Research Boot on the Policymakers' Foot: Can Participatory Approaches Change the Relationship between Policymakers and Evaluation?" in *Evidence and Evaluation in Social Policy*, edited by I. Greener and G. Bent. London, UK: Wiley Blackwell. P. 142.
[453] Partridge, E. 2011. "Caught in the Same Frame? The Language of Evidence-based Policy in Debates about the Australian Government 'Intervention' into Northern Territory Aboriginal Communities." in *Evidence And Evaluation in Social Policy*, edited by I. Greener and B. Greve. London, UK: Wiley Blackwell.
[454] Soss, J., and D. P. Moynihan. 2014. "Policy Feedback and the Politics of Administration." *Public Administration Review* 74(3):320-32. P. 322.
[455] Rist, R. C., M. H. Boily, and F. Martin. 2011. *Influencing Change. Building Evaluation Capacity to Strengthen Governance*. Washington, D.C., US: The International Bank for Reconstruction and Development/The World Bank.
[456] Heider, C. 2011. "A Conceptual Framework for Developing Evaluation Capacities: Building on Good Practice." in *Influencing Change. Building Evaluation Capacity to Strengthen Governance*, edited by R. C. Rist, M. H. Boily, and F. Martin. Washington, D.C., US: The World Bank.

values and attitudes.'[457] The outcomes of evaluation can also be perceived in its relationship to accountability.[458]

Evaluation can generate information about the effectiveness of public policy, while it can also identify new public problems and determine their magnitude.[459] It might also strengthen budgetary processes and provide evidence about the appropriate allocation and use of public funds.[460] Moreover, evaluation is especially useful in the construction of arguments for the justification and legitimacy of public action.[461]

A potential problem that evaluation and accountability share is what Bovens calls *the problem of many eyes and many hands.* This refers to the fact that 'public institutions are accountable to a plethora of different forums',[462] all of which 'apply a different set of criteria.'[463] Public agencies systematically face a wide range of requests, which implies administrative efforts that can affect the development of the primary activities of the organisation. Although this problem was developed within the framework of accountability, it can also apply to evaluation. Public agencies receive continual requests for evaluating different features of their programmes and it can become a very time-consuming activity. In the long-term, this can turn evaluation into a meaningless activity, the only purpose of which is to fulfil a legal obligation. This also finds echo in Flinders' ideas about the integrity/efficiency dilemma of

---

[457] McDonald, B., P. Rogers, and B. Kefford. 2003. "Teaching People to Fish? Building the Evaluation Capability of Public Sector Organizations." *Evaluation* 9(1):9-29. P. 11.

[458] For further reference see: Bovens, M., R.E. Goodin, and T. Schillemans (Eds.). 2014. *The Oxford Handbook of Public Accountability*. Oxford, UK: Oxford University Press.

[459] Hoefer, R. 2000. "Accountability in Action?: Program Evaluation in Nonprofit Human Service Agencies." *Nonprofit Management and Leadership* 11(2):167-77.

[460] Cunill-Grau, N., and S. M. Ospina. 2012. "Performance measurement and evaluation systems: Institutionalizing accountability for governmental results in Latin America." *New Directions for Evaluation* 2012(134):77-91.

[461] Gris Legorreta, P. C. 2011. "La evaluación de políticas públicas como instrumento para la rendición de cuentas." *Trimestre Fiscal* 99:89-103.

[462] Bovens classifies the different forums as follows: Political (Elected Representatives, Political Parties, Voters, Media); Legal (Courts); Administrative (Auditors, Inspectors and Controllers); Professional (Professional peers) and Social: (Interest Groups, Charities and Other Stakeholders). Bovens, M. 2007. "Analysing and Assessing Accountability: A Conceptual Framework." *European Law Journal* 13(4):447-68.

[463] Ibid. P. 455.

accountability.[464] *The problem of many hands* has a similar effect:

> Policies pass through many hands before they are actually put into effect […] decrees and decisions are often made in committees and cross a number of desks before they (often at different stages and at different levels) are implemented.[465]

The participation of many actors during implementation impedes identifying those who are accountable for that policy. In evaluation, when findings show poor performance, it is also difficult to determine what exactly went wrong during the process. Evaluation and accountability can be complementary activities, but it is important to consider the possible limits of evaluation and the need for developing effective institutional arrangements that can maximise this relationship. Here, Hood's reflections upon accountability and blame-avoidance 'as opposite sides of the same coin' make sense.[466] Evaluation can be both used for accountability purposes, e.g. by generating evidence about the accomplishment of goals as well as for blame-avoidance, e.g. by discretionarily disseminating only positive findings. This duality will depend on the political intention underpinning evaluation.

The theoretical review presented here guided by the thematic framework has elucidated the discussions considered most relevant for understanding the relationship between evaluation and politics. Taking this into consideration, it is useful to make some important remarks:

1. The recognition of the political nature of evaluation and the way in which different stakeholders can influence this process.
2. Evaluation is no longer a purely governmental task; it now implies the participation of new actors with divergent and sometimes conflicting interests.
3. The institutional framework for evaluation can restrain the political influence of stakeholders. There are important elements of the process of evaluation that can be regulated through the establishment of formal rules, e.g. commissioning, but this does not completely eliminate spaces for stakeholders to act discretionarily.

The acknowledgement of the political nature of evaluation favours being aware

---

[464] Flinders, M. 2011. "Daring to be a Daniel: The Pathology of Politicized Accountability in a Monitory Democracy." *Administration & Society* 43(5):595-619.
[465] Bovens, M. 2007. "Analysing and Assessing Accountability: A Conceptual Framework." *European Law Journal* 13(4):447-68. P. 457.
[466] Hood, C. 2014. "Accountability and Blame-Avoidance." in *The Oxford Handbook of Public Accountability*, edited by M. Bovens, R. E. Goodin, and T. Schillemans. Oxford, UK. P. 608.

of the potential negative effects associated, e.g. manipulation of findings, and fostering those political elements that can be beneficial, e.g. wider dissemination of findings. The next section refers back to the governance model in order to relate the discussions presented here.

2.3.3 The Hidden Politics of Evaluation in the Context of Governance

Governance has represented a substantial change in the way policymaking occurs. One of the most evident modifications has been the establishment of networks of stakeholders from different spheres that engage in the dynamics of cooperation and negotiation of interests.[467] Here the role of evaluation has been transformed beyond its capacity of generating evidence about policy. Its instrumental dimension has moved towards a new one, more related to the legitimisation of public action derived from social transformations and the rise of political demands.[468] This can be explained by the acknowledgement of the new needs of stakeholders. In a democratic context, government can no longer establish unilateral relationships; it is now compelled to negotiate and to develop cooperation networks in order to advance its agenda.

The multiplicity of interests involved in public policy demands the adoption of tools to reduce uncertainty/risk that can provide stakeholders with more elements to act. It becomes indispensable to identify 'who key policy players are, their historical roles in framing policy debates, their current policy objectives, and the consequences of these policy players' influence on evaluation practice.'[469]

This perspective highlights the significance of evaluation as a trigger for public debate. Rather than discussion based on values and individual partisan opinions, evaluation provides inputs for building more robust arguments, which are not exempted from these values and beliefs but are the result of a credible and independent process. This underlines the relevance of the legitimacy of

---

[467] Stame, N. 2006. "Governance, Democracy and Evaluation." *Evaluation* 12(1):7-16.
[468] House, E. R. Ibid."Democracy and Evaluation."119-27.
[469] Rutkowski, D., and J. Sparks. 2014. "The new scalar politics of evaluation: An emerging governance role for evaluation." Ibid.20(4):492-508. P. 504.

evaluation.[470] This explains why some evaluation approaches have become more appreciated by governments. For instance, RCTs and participatory approaches are now recurrent models for the evaluation of public policy.[471] This is not casual. RCTs are perceived as robust tools for identifying the effects of a public intervention,[472] while participatory approaches are seen as a more democratic and transparent means to involve stakeholders.[473]

The institutional framework gains importance as rules and procedures are required to enhance the role of evaluation in governance. This means setting out minimal conditions for its application.[474] A more proactive response from stakeholders is needed and responsibilities for conducting evaluations are now spread among them. Negotiating interests must be now acknowledged as a routine practice in evaluation, although the conduction of this process should not rely only on evaluators. While it is true that their role as advocates is important,[475] empowering other stakeholders is also relevant and the consequences of this must be studied in more detail.[476] Evaluation can also represent a driver for change[477] that is required in the context of the transformation of the state. The discussion about depoliticisation promotes reflecting about how transferring policy issues to other political arenas should be evaluated and, more importantly, how these evaluations should be linked to accountability.

From an overall perspective, the role of evaluation in governance appears still to be at a stage of definition. This incremental transformation will generate different issues of discussion that will need to be developed in the light of the

---

[470] Chelimsky, E. 1987. "What Have We Learned about the Politics of Program Evaluation?" *Educational Evaluation and Policy Analysis* 9(3):199-213.
[471] Pearce, W., and S. Raman. 2014. "The new randomised controlled trials (RCT) movement in public policy: challenges of epistemic governance." *Policy Sciences* 47(4):387-402.
[472] Gertler, P. J., S. Martínez, P. Premand, L. B. Rawlings, and C. M. J. Vermeesch. 2011. "La evaluación de impacto en la práctica." Washington, D.C., US: Banco Internacional de Reconstrucción y Fomento/Banco Mundial.
[473] House, E. R. 2006. "Democracy and Evaluation." *Evaluation* 12(1):119-27.
[474] Picciotto, R. 2015. "Democratic evaluation for the 21st century." Ibid.21(2):150-66.
[475] Mohan, R. 2014. "Evaluator Advocacy: It Is All in a Day's Work." *American Journal of Evaluation* 35(3):1-7.
[476] Gildemyn, M. Ibid."Understanding the Influence of Independent Civil Society Monitoring and Evaluation at the District Level: A Case Study of Ghana." (4):507-24.
[477] Molas-Gallart, J. 2012. "Research Governance and the Role of Evaluation: A Comparative Study." Ibid.33:583-98.

political context in which evaluation takes place. The aim of this chapter has been to provide the theoretical and conceptual foundations for this thesis. It has engaged with the question of *what* this thesis is really about. The next chapter focuses on *how* the thematic framework developed in this chapter will be utilised during this research.

Chapter 3. Methodology

This chapter concerns the way in which the hidden politics of evaluation has been studied. The reliability of research depends on the methodology used for the collection, systematisation and analysis of the data that led to the formulation of conclusions about an object of study. Methods are 'techniques recognised by most social scientists as being appropriate for the creation, collection, coding, organisation and analysis of data.'[478] In the social sciences, 'the scientific study of human behaviour'[479] is carried out through the use of specific instruments that can provide the best possible means for grasping the complexity of social reality. The conceptualisation of social problems is built around the different perspectives adopted for this purpose. In this field, there is no absolute nor undebatable truth; only different interpretations about an object of study. The comprehension of social reality varies also in relation to the scope adopted and the methods selected. The epistemological debate about how knowledge is constructed requires looking at multiple ideological positions about the interpretation of reality.[480] Such discussions would exceed the scope of this research. However, the strength of a research of this nature is based on the robustness of the method applied and on its capacity to simplify the reality of the object of study. Consequently, the knowledge generated about the hidden politics of evaluation needs to be supported by robust instruments to explain how the different elements analysed have been interpreted.

The conditions in which evaluation occurs require a careful study of the contextual factors affecting this process. The scope or methodological approaches selected affect the interpretation of how politics and evaluation relate. This chapter sets out the methodology used in this research and the rationale behind the selection of the approach and the analytical instruments. It is divided into three parts. The first section presents an overview of the

---

[478] Perri 6, and C. Bellamy. 2012. *Principles of Methodology: Research Design in Social Science*. Los Angeles: SAGE. P. 9.

[479] Punch, K. F. 2014. *Introduction to Social Research: Quantitative and Qualitative Approaches*: California, US. P. 9.

[480] For further reference about the different perspectives on social science research see May, T. 2011. "Social Research: Issues, Methods and Research." California, US: Open University Press.

characteristics of qualitative research (the approach selected) and the features of the case study (the analytical instrument). This section explains the rationale for these choices and the motivations for the selection of the case studies (UK and Mexico). The second part describes the process undertaken for the development of this thesis. Finally, the limitations of the analysis and the methodological challenges faced during the process are addressed.

3.1 The Qualitative Research Approach and the Case Study Method

Different perspectives and several methods exist for the study of the social sciences. According to Babbie, social research has three purposes: exploration, description, and explanation, which guide and define methodological decisions such as the scope and time framework.[481] Discussions have mostly revolved around which are the most suitable methods for social research. This debate has generated an important body of knowledge about the advantages and disadvantages of different methods in various contexts and under multiple conditions. Therefore, as Creswell states, quantitative and qualitative approaches should not be seen as dichotomies, but as a continuum.[482] This means looking at methods as a wide range of tools with multiple strengths and limitations. Despite no consensus having been reached about the superiority of a particular method for analysing social problems,[483] the core element of this decision relates to the nature of the data to be analysed.[484]

In recent decades, for example, researchers have attempted to address more complex objects of study through the use of mixed methods.[485] Combining

---

[481] Babbie, E. 2004. *The Practice of Social Research*. London, UK: Thomson/Wadsworth.

[482] Creswell, J. W. 1994. *Research Design: Qualitative and Quantitative Approaches*. California, US.: SAGE. P. 3.

[483] See Bryman, A. 1984. "The Debate about Quantitative and Qualitative Research: A Question of Method or Epistemology?" *The British Journal of Sociology* 35(1):75-92, Lawrence Neuman, W. 2014. *Basics of Social Research: Qualitative & Quantitative Approaches*. London, UK: Pearson Education Limited, Patten, M. L. 2007. *Understanding Research Methods: An Overview of the Essentials*. California, US: Pyrczak Publishing, Sale, J. E M., L. H. Lohfeld, and K. Brazil. 2002. "Revisiting the Quantitative-Qualitative Debate: Implications for Mixed-Methods Research." *Quality and Quantity* 36(1):43-53.

[484] Punch, K. F. 2014. *Introduction to Social Research: Quantitative and Qualitative Approaches*: California, US.

[485] According to the SAGE Dictionary of Social Research Methods, a mixed method research implies 'the combined use of both quantitative and qualitative methodologies within the same

methods can help overcoming the possible weaknesses found in each method individually.[486] However, as Bryman states: 'Research methods are ineluctably rooted in epistemological and ontological commitments.'[487] Methodological consistency represents therefore a key element.[488] Both quantitative and qualitative approaches serve different purposes according to the nature of the object of study and the expectations of the research.[489] The selection of the methodological approach transcends the mere collection and analysis of data; it entails reflecting on the nature of the object and the expected outcomes. Also, it involves acknowledging the characteristics of the research question and weighing what the different approaches offer for addressing it.

In political science, addressing research questions that include variables difficult to quantify, e.g. power or cooperation, is a frequent situation. As Burnham et al. observe:

> Unlike the situation in natural science, where it is often argued, there are standardized and accepted measures, social scientists are more aware that there is a range of techniques that they can choose from and a particular problem does not necessarily suggest a particular method [...] there is no correct method, even for a particular type of research.[490]

An important concern within this field is the need for expanding the range of methods in the aim of promoting more diversity in research.[491] In this context, qualitative methods have been widely utilised for understanding more carefully the diversity of phenomena that occur in the social sciences.[492] The qualitative approach claims to have a better capacity ...

---

study in order to address a single research question.' Jupp, V. 2006. *The SAGE Dictionary of Social Research Methods*. London, UK: SAGE. P. 179.

[486] Greene, J. C. 2007. *Mixed Methods in Social Inquiry*. California, US: Jossey-Bass, Johnson, R. B., and A. J. Onwuegbuzie. 2004. "Mixed Methods Research: A Research Paradigm Whose Time Has Come." *Educational Researcher* 33(7):14-26.

[487] Bryman, A. 2012. *Social Research Methods*. Oxford, UK: Oxford University Press. P. 629.

[488] O'Leary, Z. 2004. *The Essential Guide to Doing Research*. London, UK: SAGE.

[489] Mahoney and Goertz compared the quantitative and qualitative approaches. See Mahoney, J., and G. Goertz. 2006. "A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research." *Political Analysis* 14(3):227-49.

[490] Burnham, P., K. Gilland Lutz, W. Grant, and Z. Layton-Henry. 2008. *Research Methods in Politics*. Basingstoke, UK: Palgrave Macmillan. Pp. 2-3.

[491] Kinder, D. R., and T. R. Palfrey. 1993. *Experimental Foundations of Political Science*: University of Michigan Press.

[492] For example: Ritz, A. 2011. "Attraction to Public Policy-Making: A Qualitative Inquiry into Improvements in PSM Measurement." *Public Administration* 89(3):1128-47, Van Der Heijden, J. 2014. "The Role of Government in Voluntary Environmental Programmes: A Fuzzy Set Qualitative Comparative Analysis." Ibid.

... to constitute compelling arguments about *how things work in particular contexts* [...] the qualitative habit of intimately connecting context with explanation means that qualitative research is capable of producing very well-founded *cross-contextual generalities,* rather than aspiring to more flimsy de-contextual versions.[493]

This approach facilitates the interpretation of multifactorial phenomena.[494] The use of qualitative research in certain policy fields, i.e. social policy, has been prominent because it is 'underpinned by the persistent requirement [...] to understand complex behaviours, needs, systems and cultures.'[495] As a result, it provides an analytical framework for reaching more specific levels of comprehension about the interactions that take place among individuals and that might affect the outcomes of a process, e.g. public policy. In terms of comparative politics, for example, Mahoney suggests that qualitative research 'offers [...] various tools for framing research questions and formulating testable hypotheses.'[496] Specifically in policymaking, Pierson states that in contrast to quantitative methods the use of a qualitative approach is desirable because it provides a broader perspective of its functioning by 'looking at the long haul.'[497] Ritz agrees by stating that qualitative research 'allows a more in-depth and inner perspective of the object of interest than quantitative research.'[498] This shows the potential of the qualitative approach for the study of political issues derived from its capacity for a more detailed analysis, which can capture the relationships that take place in a particular context.

The identification of the potential uses of the qualitative approach for the core argument of this thesis is based precisely on the complexity associated with politics. The scope of analysis needs to allow not only the study of the relationships established as a consequence of evaluation, but also how these are affected by rules and institutions. Although methodological choices can always

---

[493] Mason, J. 2002. *Qualitative Researching*. London, UK: Sage. P. 1.
[494] Goertz, G. 2006. "Introduction to the Special Issue "Causal Complexity and Qualitative Methods"." *Political Analysis* 14(3):223-26.
[495] Huberman, A. M., and M. B. Miles. 2002. *The qualitative researcher's companion*. London, UK: SAGE. P. 305.
[496] Mahoney, J. 2007. "Qualitative Methodology and Comparative Politics." *Comparative Political Studies* 40(2):122-44. P. 124.
[497] Pierson, P. Ibid."The Costs of Marginalization: Qualitative Methods in the Study of American Politics."146-69.P. 155.
[498] Ritz, A. 2011. "Attraction to Public Policy-Making: A Qualitative Inquiry into Improvements in PSM Measurement." *Public Administration* 89(3):1128-47. P. 1133.

be debatable, the strength of research relies on recognising its suitability for the particular object of study and the context in which it occurs. The next section discusses the main characteristics of the qualitative approach that make explicit its pertinence for studying the hidden politics of evaluation.

3.1.1 General Characteristics of the Qualitative Approach

The qualitative approach can be defined as 'research that investigates aspects of social life which are not amenable to quantitative measures.'[499] According to Lawrence, this approach has the capacity to analyse the complex social reality, as it "refines 'working ideas' during the data collection and analysis process [to] form coherent theoretical definitions as you struggle to 'make sense' or organize the data and the preliminary ideas."[500] The logic that underpins the research of this nature is built from different theories and methods.[501] According to Bryman, qualitative research presents three core characteristics:

1. an inductive view of the relationship between theory and research [...]
2. an epistemological position described as interpretivist [...] the stress in on the understanding of the social world through an examination of the interpretation of that world by its participants; and,
3. an ontological position described as constructionist, which implies that social properties are outcomes of the interactions between individuals, rather than phenomena 'out there' and separate from those involved in its construction.[502]

Punch agrees by underlining that qualitative research denotes that ...

> research [...] is a human construction, framed and presented within a particular set of discourses (and sometimes ideologies), and conducted in a social context with certain sorts of social arrangements, involving especially funding, cognitive authority and power.[503]

Due to its characteristics, the qualitative approach does not fit into an inflexible research design. Therefore, the robustness of the methodology demands being explicit about the different stages of the process and the activities undertaken

---

[499] Jupp, V. 2006. *The SAGE Dictionary of Social Research Methods*. London, UK: SAGE. P. 248.
[500] Lawrence Neuman, W. 2014. *Basics of Social Research: Qualitative & Quantitative Approaches*. London, UK: Pearson Education Limited. P. 139.
[501] Flick, U. 2014. *An Introduction to Qualitative Research*. California, US: SAGE.
[502] Bryman, A. 2012. *Social Research Methods*. Oxford, UK: Oxford University Press. P. 380.
[503] Punch, K. F. 2014. *Introduction to Social Research: Quantitative and Qualitative Approaches*: California, US. P. 117.

for obtaining the findings. [504] Qualitative research, therefore …

> … covers a wide range of approaches, but by definition, none of these
> approaches relies on numerical measurements […] such work has tended
> to focus on one or a small number of cases, to use intensive interviews or
> depth analysis of historical materials, to be discursive in method, and to be
> concerned with a rounded or comprehensive account of some event or
> unit.[505]

It aims to make inferences from the analysis of data obtained through multiple instruments, e.g. case studies, interviews, etc. However, these are subject to the particular conditions of the object of study and cannot be translated into a different context. It imposes creating mechanisms and guidelines appropriate for the reality to study.[506] The way in which qualitative research is conducted cannot be conceived as a sequential completion of stages.[507] Instead, it occurs according to the different activities planned, developed across time, and is subject to the conditions existing in the context of the research. The general conceptualisation of qualitative research suggests that it is a continuous process that involves the consecution of activities—in occasions overlapped—with which outputs eventually lead to the identification of findings that shape the course of the research. The key element here is the capacity to interpret the reality through the evidence obtained.

The way in which data is produced is also relevant. For example, using quantitative methods for this purpose represents only one alternative for causal interpretation. Nonetheless, as Denzin states:

> Causality as a type of narrative is only one form of interpretation […]
> autoethnographic, action-based and other forms of narrative
> representation are equally powerful methods and strategies of analysis and
> interpretation.[508]

In this particular research, the analysis of the hidden politics of evaluation does not aim to identify causes but to understand how this phenomenon manifests

---

[504] Mason, J. 2002. *Qualitative Researching*. London, UK: Sage.
[505] King, G., R. O. Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. New Jersey, US: Princeton University Press*. P. 4.
[506] Denzin, N. K. 2009. "The elephant in the living room: or extending the conversation about the politics of evidence." *Qualitative Research* 9(2):139-60.
[507] Flick, U. 2014. *An Introduction to Qualitative Research*. California, US: SAGE.
[508] Denzin, N. K. 2009. "The elephant in the living room: or extending the conversation about the politics of evidence." *Qualitative Research* 9(2):139-60. P. 146.

and relates to the institutional framework. This distinction matters because the focus would be on the nexus between politics and evaluation and its implications. Knowing why politics affects evaluation is a sterile effort if the consequences of this are not studied. There is an important range of tools for collecting, systematising, and analysing qualitative data.[509] The suitability of these tools for the research depends on several reasons. For example, the nature of the object of study, the access to it, and the temporality are just some variables to consider in the selection of a qualitative method. The selection of a particular method or a combination will be based on the criteria of the researcher and the prioritisation of those characteristics that are considered essential for the research.

The nature of the qualitative approach and its capacity to grasp the interaction between individuals is strongly related to the study of the hidden politics of evaluation in the sense that this research aims is to understand the outcomes and implications of this phenomenon and how it responds to the rules and institutions developed for evaluation.

3.1.2 The Case Study Method

Different scholars in the social sciences literature have attempted to define 'the case study'.[510] Simons, for example, conceptualises this as …

> … an in-depth exploration from multiple perspectives of the complexity and uniqueness of a particular project, policy, institution, programme or system in a 'real life' context […] The primary purpose is to generate in-depth understanding of a specific topic (as in a thesis), programme, policy institution or system to generate knowledge and/or inform policy development, professional practice and civil or community action.[511]

The case study has a predominant role in the development of new knowledge in

---

[509] Creswell, J. W. 1994. *Research Design: Qualitative and Quantitative Approaches*. California, US.: SAGE, King, G., R. O. Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. New Jersey, US: Princeton University Press, Mason, J. 2002. *Qualitative Researching*. London, UK: Sage, Punch, K. F. 2014. *Introduction to Social Research: Quantitative and Qualitative Approaches*: California, US.

[510] See Mills, A. J., G. Durepos, and E. Wiebe. 2010. "Case Study as a Methodological Approach. Encyclopedia of Case Study Research." Pp. 67-78. London, UK: SAGE, Thomas, G. 2011. *How to do your case study: A guide for students*. London, UK: SAGE, Yin, R. K. 2009. *Case Study Research: Design and Methods*. London, UK: SAGE, ibid.

[511] Simons, H. 2008. *Case Study Research in Practice*. London, UK: SAGE. P. 21.

the social sciences because of the characteristics of this method. According to Yin:

> The case study method allows investigators to retain the holistic and meaningful characteristics of real-life events –such as individual life cycles, small group behaviour, organizational and managerial processes, neighbourhood change, school performance, international relations, and the maturation of industries.[512]

In a sense, this tool facilitates abstracting a particular phenomenon from the complex context in which it occurs and offers the capacity to reach more specific levels of analysis:

> Most case studies feature descriptions that are complex, holistic, and involving a myriad of not highly isolated variables; data that are likely to be gathered at least partly by personalistic observation; and a writing style that is informal, perhaps narrative, possibly with verbatim quotation, illustration, and even allusion and metaphor [...] themes and hypotheses may be important, but they remain subordinate to the understanding of the case.[513]

It provides detailed knowledge about a particular object. Consequently, it is useful for looking at its different dimensions and for studying in detail what occurs in a specific context, by 'getting a rich picture and gaining analytical insights from it.'[514] These characteristics show the capacity of this method to disaggregate different variables and to observe an object from different angles, in contrast to other methods (Table 3.1).

---

[512] Yin, R. K. 2009. *Case Study Research: Design and Methods.* London, UK: SAGE. P. 4.

[513] Stake, R. E. 1978. "The Case Study Method in Social Inquiry." *Educational Researcher* 7(2):5-8. P. 7.

[514] Thomas, G. 2011. *How to do your case study: A guide for students.* London, UK: SAGE. P. 23.

Table 3. 1 A Comparison of the Case Study with Other Forms of Inquiry

	Case Study	Experiment	Survey
Investigates…	one case or a small number of cases	a relatively large number of cases	a relatively large number of cases
Data collected and analysed about…	a large number of features of each case	a small number of features of each case	a large number of features of each case
Study of…	naturally occurring cases where the aim is not to control the variables	cases where the aim is to control the important variables	naturally occurring cases selected to maximise the sample's representativeness of a wider population
Quantification of data…	is not a priority	is a priority	is a priority
Using…	many methods and sources of data	one method	one method
Aiming to…	look at relationships and processes	look at causation	look for generalisation

Source: Thomas, G., 2011. *How to do your case study: A guide for students.* SAGE. Pp. 10-11.

The case study is considered an important methodological tool for the analysis of complex topics.[515] Within the political sciences, Gerring recognises that this method is useful when 'insight into causal mechanisms is more important than insight into causal effects.'[516] This resonates with the core theme of this thesis, as knowing which are the elements of evaluation that are affected by politics, in the context of a particular institutional design, requires a more detailed study than explaining why politics affects evaluation. In the case of evaluation research, the use of the case study is pertinent due to its political nature. The inclusion of multiple perspectives and the interpretation of the values involved can be captured through this instrument.[517] According to Yin, evaluation research can benefit from this method for two main reasons:

> First, the ability directly to incorporate an investigation of the context satisfies an evaluation's need to monitor and assess both the intervention and the implementation process. Second, the case study is not limited to

---

[515] For further reference about the different kinds of case studies used in social and political research see: Moses, J. W., and T. L. Knutsen. 2012. *Ways of Knowing: Competing Methodologies in Social and Political Research.* New York, US: Palgrave Macmillan.

[516] Gerring, J. 2004. "What Is a Case Study and What Is It Good for?" *American Political Science Review* 98(2):341-54. P. 352.

[517] Simons, H. 2008. *Case Study Research in Practice.* London, UK: SAGE.

either quantitative or qualitative data, but can incorporate both varieties of evidence [...] the case study method can serve evaluation needs directly by being able to assess outcomes and to test hypotheses.[518]

In line with this, Keen and Packwood observe that case studies in health policy are useful when looking at the success or failure of a particular intervention when there are contextual factors to consider.[519] This agrees with Yin's argument about its capacity to test hypotheses.[520] Therefore, this method offers analytical lenses for comprehending a phenomenon occurring in a particular context.[521] As a result, some of its strengths relate to its capacity to capture the experiences of the actors involved and its focus on causal relationships between variables.[522] In line with the idea of the institutional framework, it can also 'be useful in capturing the emergent and immanent properties of life in organizations.'[523]

Despite the advantages recognised in this method, it is also important to acknowledge its limitations. Perhaps the main criticism of this method refers to the impossibility of generalisation of the conclusions because of validity reasons,[524] as 'in most circumstances the individual cases are not sufficiently representative.'[525] Nonetheless, as Flyvbjerg recognises, 'a purely descriptive, phenomenological case study without any attempt to generalize can certainly be

---

[518] Yin, R. K. 1992. "The Case Study Method as a Tool for Doing Evaluation." *Current Sociology* 40(1):121-37. P. 124.

[519] Keen, J., and T. Packwood. 1995. "Case Study Evaluation." *BMJ: British Medical Journal* 311(7002):444-46.

[520] The use of this method in evaluation has been analysed and promoted by organisations such the United States General Accounting Office and The World Bank. See: Morra, L. G., and A. C. Friedlander. 1999. "Case Study Evaluations." edited by The World Bank, United States General Accounting Office. 1990. "Case Study Evaluations."

[521] Scriven, M. 1991. *Evaluation Thesaurus*. London, UK: SAGE.

[522] Hodkinson, P., and H. Hodkinson. 2001. "The Strengths and Limitations of Case Study Research." in *Making an Impact on Policy and Practice*. Cambridge, UK.

[523]Noor, K.B.M. 2008. "Case Study: A Strategic Research Methodology." *American Journal of Applied Sciences* 5(11):1602-04. P. 1603.

[524] The quality of social research has been appraised through four major tests: '*construct validity:* identifying correct operational measures for the concepts being studied; *internal validity* (for explanatory or causal studies only and not for descriptive or exploratory studies): seeking to establish a causal relationship, whereby certain conditions are believed to lead to other conditions, as distinguished from spurious relationships; *external validity*: defining the domain to which a study's findings can be generalized; and *reliability:* demonstrating that the operations of a study –such as the data collection procedures- can be repeated with the same results.' Yin, R. K. 2009. *Case Study Research: Design and Methods*. London, UK: SAGE. P. 40.

[525] Jupp, V. 2006. *The SAGE Dictionary of Social Research Methods*. London, UK: SAGE. P. 20.

of value.'[526]

Amaratunga and Baldry add that this method 'suffers from a lack of rigour and an excess of bias.'[527] This can be explained by the fact that the findings produced through a case study are mostly contextual. But it might be precisely this feature that permits valuing different explanations for a phenomenon.[528] As Gerring states, political science case studies are useful 'when propositional depth is prized over breadth and boundedness, [and] when (internal) case comparability is given precedence over (external) case representativeness.'[529] These characteristics might help overcoming some of the weakness identified in the method.

This general overview of the qualitative approach and the case study aid to make explicit the methodology applied in this research. The qualitative approach and the case study are the methods on which this research relies. Documentary analysis and semi-structured interviews were the instruments used for the collection and analysis of data. Considering the characteristics of the method exposed here, the next section explains the rationale behind this methodological selection.

3.1.3 Rationale for the Selection of the Methodology

The study of the political nature of evaluation and its implications requires looking at particular contexts in which this phenomenon occurs to identify the different areas of the process that are more vulnerable to this influence. The purpose of the selected methodological approach is to provide the analytical tools required for grasping those variables useful for uncovering the hidden politics of evaluation. As discussed, the qualitative approach offers a suitable methodological framework for analysing the relationships that are established during evaluation and the constraints that the institutional framework imposes

---

[526] Flyvbjerg, B. 2006. "Five Misunderstandings About Case-Study Research." *Qualitative Inquiry* 12(2):219-45. P. 227.
[527] Dilanthi, A., and D. Baldry. 2001. "Case study methodology as a means of theory building: performance measurement in facilities management organisations." *Work Study* 50(3):95-105. P. 100.
[528] See also Meyer, C. B. 2001. "A Case in Case Study Methodology." *Field Methods* 13(4):329-52.
[529] Gerring, J. 2004. "What Is a Case Study and What Is It Good for?" *American Political Science Review* 98(2):341-54. P. 352.

to these interactions. Considering that the hidden politics of evaluation do not manifest in the same way in every context, the case study method helps recognising those contextual factors that might explain variances in the behaviour of stakeholders in this process. The selection of this methodology can be explained by the following reasons:

1. Most of the elements analysed in this research are not quantifiable, e.g. interaction between stakeholders. Consequently, these cannot be measured through a quantitative analysis. Documentary analysis and interviews are more appropriate analytical tools to capture the complexity of the relationships that are established during evaluation.
2. The qualitative approach is helpful for understanding the hidden politics of evaluation, as it is an intangible component that interacts with different elements such as power, asymmetric information, and economic incentives. Moreover, the case study allows looking at variables identified by actors involved in the process and contrasting different perspectives that can be triangulated.[530]
3. A case study comparison helps observing how a set of variables behave in two different contexts. This approach enables the researcher 'to test general hypotheses about politics'[531] by identifying similarities and differences between cases that can sustain the findings obtained.[532] The institutional architecture, as the object of study, allows looking at the way in which these rules and procedures affect the elements of the thematic framework proposed.
4. A case study is directed to contribute to the theory by analysing a particular situation in depth. More than testing a theory (which will require a different methodological treatment), this thesis aims to understand how different relations are built within evaluation stakeholders and the implications for the process.
5. The thematic framework developed for the understanding of the hidden politics of evaluation is strongly related to the conditions of the institutional framework in which evaluation takes place.[533] By focusing on the evaluation of health policy in both countries and how rules and procedures affect this process, it is possible to 'highlight cross-national differences, and their causes and consequences, in an effective fashion.'[534]
6. The differences in the institutional framework for evaluation found in the case studies provide new elements for analysing how politics and evaluation interact.

These reasons highlight two main characteristics of this research. The first one

---

[530] The concept of triangulation in social science research refers to the process of combining methods (e.g. interviews and documentary analysis) in order to obtain more reliable findings, for example, by contrasting different views about a same issue. See: Jick, T. D. 1979. "Mixing Qualitative and Quantitative Methods: Triangulation in Action." *Administrative Science Quarterly* 24(4):602-11.

[531] Hague, R., M. Harrop, and S. Breslin. 1992. *Comparative Government and Politics*. London, UK: The McMillan Press Ltd. P. 27.

[532] See Kaarbo, J., and R. K. Beasley. 1999. "A Practical Guide to the Comparative Case Study Method in Political Psychology." *Political Psychology* 20(2):369-91.

[533] Pollitt and Bouckaert's comparative analysis about the adoption of public management reforms in different countries is a key text for understanding for example how the politico-administrative regime affects this process. Pollitt, C., and G. Bouckaert. 2000. *Public Management Reform: A Comparative Analysis*. Oxford, UK: Oxford University Press.

[534] Hague, R., M. Harrop, and S. Breslin. 1992. *Comparative Government and Politics*. London, UK: The McMillan Press Ltd. P. 36.

is the intangibility of the object of study (the hidden politics of evaluation) and the need for its interpretation. The other refers to the existence of a common explanatory variable (the institutional framework) that requires developing instruments for the comparability of cases. The analysis of the institutional framework is guided by the elements of the thematic framework proposed to identify the set of rules, procedures, and actors involved in evaluation and to facilitate the comparison between countries. This methodological approach provides the analytical tools for conducting the research, which process is explained in the next section.

3.2 Research Methodology

The description of the process undertaken for the collection, systematisation, and analysis of the data is crucial for identifying how the conclusions were obtained and how this information can be interpreted. The entire research process implied reflecting continuously on the main and secondary research questions that guide this thesis and how these can be operationalised. The process through which the elements of discussion were obtained is just as relevant as the findings.[535] The hidden politics of evaluation represents an interesting opportunity for the development of research models that can be adapted to its complexity and that can give new insights about how the institutional framework affects evaluation and vice versa. This section explains how this research was conducted and the methodological grounds on which it is based. The research design and process are described in detail, as well as the rationale for the selection of the case studies.

3.2.1 Research Design

The research design was developed around the understanding of the relationship between politics and evaluation. In line with this, the core question was disaggregated into more specific elements to establish a clearer idea about the process of knowledge construction (see Chapter 1, Section 1.1). These

---

[535] An analytical research 'seeks to explain why things are as they are, or how they came to be as they are, by a process of analysis and explanation.' In line with this conceptualization, this research aims to explain how politics and evaluation interact by looking at a set of variables related to the assessment process in which politics manifest. Brewer, R. 2007. *Your PhD Thesis. How to plan, draft, revise and edit your thesis*. Abergele, UK: Studymates Limited. P. 17.

questions concern different elements of evaluation, of which analysis can reveal its connection to politics. For example: *Does the evaluation policy produce relevant and convenient information that can be used for policymaking?* This allows exposing the association of key elements of evaluation with the hidden politics of evaluation, and later, those with the empirical findings, e.g. *How are evaluation findings used in Mexico?*

The flexible structure of the qualitative approach favours the incremental building of ideas and the development of different hypotheses during the course of the research. This process is driven by the identification and analysis of data to expose how the hidden politics of evaluation manifests. The richness of the research relies on the process of obtaining the perceptions and experiences of the stakeholders of evaluation, considering the complexity and intangibility of this object of study. This is consistent with the idea that the different rules established for evaluation will produce a different dynamic between stakeholders.

The methodology applied makes explicit the connection between the variables associated with the hidden politics of evaluation and the particular characteristics of the context in which evaluation occurs in the UK and Mexico. For instance, the set of rules for the performance of evaluations can provide important notions about the level of influence that evaluands can have in a study and how this affects its outcomes. The utilisation of evaluation outputs can also reveal the extent to which evaluation has permeated the organisational culture and if it has transcended its managerial role to become a political instrument. The thematic framework developed about the hidden politics of evaluation is also linked to the research design. The institutional framework, as a core element of discussion in this thesis, is relevant for acknowledging how rules determine the way in which decisions are made and how stakeholders behave under these conditions. The elements of this framework (purpose, resources, process, outputs, and outcomes) demand a detailed understanding of the perceptions and experiences of the stakeholders involved in evaluation (Table 3.2).

Table 3.2 Research Design

Research Questions	Data Sources and Methods	Justification	Chapter of the thesis in which the question is addressed
1. Are there efficient institutional frameworks that assure an effective assessment of public policy?	Documentary analysis	The identification of rules and procedures established for the performance of evaluation/audit in both countries can provide a clearer understanding of how the behaviour of stakeholders is shaped and to what extent it is possible for them to pursue their interests.	Chapters 4 and 5
2. Are the evaluation policies reliable? Are there robust instruments to assess the effectiveness of public policy?	Documentary analysis Evaluators / Programme operators interviews	The identification of the reliability of evaluation policy has to do with the robustness of the methods selected which can be known by the documentary evidence produced in both countries; however, it also incorporates an perception component that can only be seen through the experiences and empirical knowledge of programme operators who are actively engaged in the implementation of evaluation policies.	Chapters 4 and 5
3. Does the evaluation policy, in both countries, produce in a systematic way relevant and convenient information that can be used for the policy-making process?	Documentary analysis Programme operators/ Evaluators/ Politicians interviews	The outputs of evaluation can be used for different purposes; the way in which evaluation becomes an input for the construction of political arguments is a phenomenon that can only be perceived from the experience of the stakeholders involved. Nonetheless, through documentary analysis it is possible to observe how policymakers use the information derived from evaluation/audit.	Chapter 6
4. When the outputs of evaluation are used, is there any effect on the improvement of policy-making?	Programme operators/ Evaluators interviews	It is a highly complex task to isolate the effect of evaluation in the policymaking process as it is influenced by multiple factors. In order to obtain a general idea of the consequences that the performance of evaluation and audits have had in the broader context of governance it is important to know with more detail how stakeholders conceptualize this contribution in their own areas of influence.	Chapter 6

Table 3.2   *Continued*

Research Questions	Data Sources and Methods	Justification	Chapter of the thesis in which the question is addressed
5. Which is the role of policy evaluation in order to consider it useful for the policy-making process?	Documentary analysis Practitioners/ Academicians interviews	One of the key elements of the hidden politics of evaluation is the driver that promotes the use of evaluation/audit in the public sector. The rationale and motivation behind an evaluation policy determines in great measure the way in which this process will be conducted and its possible outputs. The documentary analysis at this stage will provide the 'official view' on the use of evaluation as a management tool, this needs to be contrasted with the views of stakeholders of evaluation who have a more informal knowledge of the drivers that guide the assessment of public policy.	Chapter 7
6. In terms of 'good policy' and 'good politics', which of these issues predominate in the evaluation agenda of both countries?	Politicians/ Practitioners interviews	The intangibility of the hidden politics of evaluation implies a subjective understanding according to the different views of stakeholders involved, therefore, the performance of in-depth interviews will allow obtaining data about the way in which these actors conceive and understand the political nature of evaluation.	Chapter 7

Source: Based on Mason, J., 2002. *Qualitative Researching.* SAGE.

3.2.2 Research Process

Although qualitative research cannot be seen as a linear process, it is possible to identify some key stages (not necessarily sequential) whose main purpose is the obtaining of empirical evidence for the development of new knowledge about the political nature of evaluation. A series of activities were performed in order to collect, systematise, and analyse data, as is described below:

1. The first stage of the process was a careful review of the literature in evaluation studies to identify those key elements of the hidden politics of evaluation considered crucial for the study of this phenomenon.
2. The literature review was the basis for the construction of a theoretical framework for the thesis based on three pillars: governance, public policy, and evaluation studies.
3. Based on the theoretical review, a thematic framework was developed for the understanding of the hidden politics of evaluation (see Chapter 2). This analytical tool aimed to play as a structured guidance for the comprehension of the political nature of evaluation.
4. An exploratory analysis of different evaluation and audit policies led to the selection of the cases to be analysed. In order to have a basis for comparison, it was decided that the universe of study would be composed of four case studies in two different countries (UK and Mexico).
5. To narrow the scope of the research a policy area was selected with an important development of evaluation and audit processes in both countries: health policy. The rationale behind this is that both in the UK and Mexico it is possible to identify a series of evaluation results susceptible to analysis relevant for reflecting upon its use. Although the way in which the government operates in both countries differs, it was considered pertinent to select two cases per country that belonged to the health sector and that have been systematically evaluated. In both countries it was considered the existence of an internal evaluation process and an external audit body (the NAO in the UK and the Supreme Audit Office (SAO) in Mexico). The selection of the cases was based on the criterion of assessment. This means that the nature and characteristics of each one of the cases was not the main focus of the research; the central issue to discuss was how politics influences evaluation/audit processes. Consequently, the cases only represent a vehicle for understanding this nexus.
6. After selecting the case studies, primary data was obtained through the performance of semi-structured interviews with stakeholders of evaluation in both countries. This led to the identification of possible participants who were open to collaborating with the research. After a preliminary exploration of the evaluation policies of both countries, potential interviewees were identified. This incorporated people from different sectors that are or had been directly involved in the evaluation/audit processes of the selected case studies, i.e. evaluators, operators.
7. A preliminary list of potential interviewees was defined and this stage was followed by a first attempt at contact (by e-mail, letter, or telephone) in which the purpose of the research and the relevance of their participation was briefly explained to them. With those participants who accepted taking part in the study a meeting for the interview was arranged. About 40 semi-structured interviews with different stakeholders of evaluation/audit were performed in

both countries. 95% of them were face-to-face interviews, with a minor number of informal chats performed over the telephone. Depending on the consent of the interviewee, some interviews were recorded. However, in all cases participants were informed that the data treatment would be made under the Chatham House Rule.[536]

8. Interviews were carried out during an approximate 2-month period in each country. Most of them took place at the interviewee's workplace and with an average length of 40 minutes. The interviews were semi-structured, based on the core issues identified in the thematic framework for the understanding of the hidden politics of evaluation.

9. Interviews were transcribed and a series of quotes related to the core issues of the hidden politics of evaluation were identified. Due to the nature of the research, the use of data analysis software, e.g. Nvivo, Atlas.ti, etc., was avoided to have a more accurate interpretation of the ideas discussed during the interviews. As politics represents the main variable for the analysis, an important part of it was observing the behaviour of interviewees which cannot be perceived by the software. To preserve the anonymity of participants, each interviewee was assigned a code-number (e.g. UK-12/MX-9) that allows only its identification according to the country of origin.[537]

10. Finally, the documentary analysis and the data collected from the interviews served as a basis for the development of a series of conclusions and recommendations about the hidden politics of evaluation.

This description offers a structured idea of the general activities that led to the formulation of conclusions of the research. The interaction with the interviewees, the analysis of the data, and the identification of new areas of study in relation to the hidden politics of evaluation represent an important asset for the generation of new knowledge that constitutes an intangible element of the research process.

3.2.3 Rationale for the Selection of the Case Studies

The research analyses the evaluation policies of the UK and Mexico by focusing on health policy. In particular, the objects of study were the Cancer Strategy and the CQC (UK case) and SICALIDAD and Health Caravans Programmes (Mexican case). The analysis of the evaluation and audit outputs incorporated the 2011–14 period. However, for the purposes of the research some references to prior evaluations were made. The comparative perspective adopted by this research recognises that the socio-economic and political differences between countries have influenced the institutional architecture for the development of evaluation.

---

[536] This rule states that 'when a meeting, or part thereof, is held under the Chatham House Rule, participants are free to use the information received, but neither the identity nor the affiliation of the speaker(s), nor that of any other participant, may be revealed.' The Royal Institute of International Affairs. 2011. "Chatham House Rule." London, UK.

[537] *Appendix C* presents a detailed list of interviewees.

Although these differences are not the focus of this research, it is important to make some remarks about this:

- The political systems operating in these countries (the UK a constitutional monarchy and Mexico a presidential regime) impact on the configuration of the Executive and the Legislative and how these powers are executed.[538]
- The level of maturity of these democracies also affects the establishment of control and accountability mechanisms and its effectiveness for the overview of public policy.[539]
- The political culture in both countries is an element that might underpin the definition of rules for the interaction within the public sector. In the British civil service, for example, values like: 'Integrity, honesty, impartiality, probity in the use of public money and putting the public interest above their own are still widely promoted and supported.'[540] *Ergo,* it is comprehensible that these actors have discretion at their disposal for performing their tasks. In contrast, in the Mexican case the adoption of control and accountability reforms, for instance, has derived from the excessive development of formal rules and procedures, derived from 'a profound mistrust in public officials'.[541]
- In terms of the reform of the public sector, it is interesting to compare the UK case, as a paradigmatic example of the implementation of NPM reforms to Mexico as a country in which efforts in this matter are relatively recent.

Considering this, the way in which institutions shape decision-making is relevant for understanding the potential differences that arise between case studies.[542] The implementation of health programmes and the operation of healthcare organisations share some common grounds that contribute to the comparability of the cases. In the first place, the cases analysed are operated by a central department and demand the participation of different actors outside this sphere. In addition, in both countries health policy is considered a priority issue in the political agenda, which gives this policy area more relevance in its study.[543] In both contexts it is possible to identify formal processes of evaluation

---

[538] See: "Political Constitution of the United Mexican States.", Benwell, R., and O. Gay. 2011. "The Separation of Powers." London, UK.

[539] This can be seen in the Global Democracy Ranking that assesses dimensions associated to the level of democracy in a country, i.e. political rights, civil liberties. During 2014, in this index the UK ranked 13th and Mexico 58th. Global Democracy Ranking. 2014. "The Democracy Ranking of the Quality of Democracy."

[540] Horton, S. 2006. "The Public Service Ethos in the British Civil Service: An Historical Institutional Analysis." *Public Policy and Administration* 21(1):32-48. P. 44.

[541] Arellano, D. 2007. "Una cuestión de confianza. Retos e instrumentos para una reforma del Estado en América Latina." *Nueva Sociedad* (210):95-109.P. 103.

[542] Weaver, K. R., and B. A. Rockman. 1993. *Do Institutions Matter? Government Capabilities in the United States and Abroad*. Washington, D.C., US: The Brookings Institution.

[543] For example, in Mexico during the 2001-2006 Presidential term one of the core elements of the governmental priorities was the universal health coverage, particularly through the implementation of the *Seguro Popular* programme, which was widely promoted by the Executive. In the UK, health policy has also been a top priority within the governmental sphere,

promoted by central government and by an external audit body. There are an important number of evaluation/audit reports that show how these processes are conducted and how outputs are used by stakeholders. The case of the CQC requires making an additional remark. The inclusion of an organisation (as opposed to a policy/programme) responds to the specific value of this body as an object of study and its potential for providing important evidence about the hidden politics of evaluation. Its creation as a merger of previous healthcare regulators shows how the driver for efficiency in the public sector has fostered the adoption of important changes in the bureaucratic structure. Moreover, the way in which the CQC has operated and its political salience are interesting elements to analyse in terms of politics.

Finally, as the institutional framework is one of the key elements of analysis, the identification of formal and informal rules for evaluation in both countries is useful for the comparison. In more specific terms, the selection of the case studies responded to the following reasons:

- Mexico constitutes one of the Latin American countries that have recently adopted performance-management measures to improve efficiency and effectiveness in the public sector. Consequently, it seems relevant to contrast it to the UK case which is among the first countries that adopted NPM measures. The rationale behind this selection is the observing of the influence of the institutional framework in terms of the maturity of the evaluation policy in relation to its political nature.
- The institutional framework for evaluation in these countries is substantially different. While in Mexico this process is highly centralised and conducted by three federal government agencies, in the UK evaluation is a process conducted in a more discretionary way and characterised by a strong audit body. It is interesting to analyse how these different institutional frameworks promote or inhibit political influence in evaluation.
- The comparison of these two cases allows identifying those contextual variables that might favour the negative influence of politics and that might limit the use of evaluation results in policymaking.

In addition to the rationale for the selection of the methodological approach, it is also necessary to discuss some of the challenges faced during the research process, addressed in the following section.

---

the Conservatives Manifesto 2010, for instance, incorporates different statements about the need for increasing health spending and for more patient-centred policies. See: Gobierno Federal. 2012. "Sexto Informe de Gobierno.", The Conservative Party. "The Conservative Manifesto 2015."

3.3 Limitations and Methodological Challenges

A sensitive topic like politics inevitably generates different reactions among people because of what it represents in terms of the exchange of power and resources. Moreover, scrutiny activities like evaluation and auditing have constantly found some 'natural' resistance due to the potential consequences that can be produced for individuals. A research that combines these elements is more likely to face some challenges that are important to consider, particularly for the conduction of the fieldwork.

3.3.1 Limitations of the Methodological Approach

As discussed, the case study is useful for the profound understanding of a particular phenomenon for which variables cannot be quantified. The study of the politics of evaluation fits into this description because of the intangible variables that can only be observed through the perception and experiences of participants. A first limitation to recognise regards the incapacity of generalisation associated with this method (see Section 3.1.2). There is an active theoretical discussion about this issue. For example, Gomm et al. distinguish between 'generalisation across and within cases.'[544] According to them, this matters in terms of the way in which researchers can claim the generalisation of their findings. Flyvbjerg adds 'that knowledge cannot be formally generalized does not mean that it cannot enter into the collective process of knowledge accumulation in a given field or in a society'.[545] The generalisation of findings, specially in this research, appears to be more useful for the identification of new areas of knowledge to explore.[546] In the particular case of the hidden politics of evaluation, for which there is little empirical evidence, findings can provide some guidance about those elements of the institutional framework that might

---

[544] Gomm, R., M. Hammersley, and P. Foster (Eds.). 2000. *Case Study Method: Key Issues, Key Texts*. London, UK: SAGE. P. 98.

[545] Flyvbjerg, B. 2006. "Five Misunderstandings About Case-Study Research." *Qualitative Inquiry* 12(2):219-45. P. 227.

[546] The literature about policy transfer can be considered a suitable framework for understanding how findings derived from case studies can be translated to other contexts for the development of public policy. For further reference see: Dolowitz, D. P., and D. Marsh. 1996. "Who Learns What from Whom: a Review of the Policy Transfer Literature." *Political Studies* 44(2):343-57, —. 2000. "Learning from Abroad: The Role of Policy Transfer in Contemporary Policy-Making." *Governance* 13(1):5-23, Evans, M., and J. Davies. 1999. "Understanding Policy Transfer: A Multi-Level, Multi-Disciplinary Perspective." *Public Administration* 77(2):361-85.

influence this phenomenon, regardless of the contextual factors that affect the process of evaluation.

The inferences made from the data collected and analysed apply only to the specific cases selected in both countries. This means that different findings could have been obtained through the study of cases in other countries or even different cases within the same context. This can also be influenced by the use of a different method, e.g. quantitative/mixed, for the collection and treatment of data. The conclusions reached in this thesis are strongly influenced by the perceptions and experiences of the stakeholders interviewed. Having a different set of interviewees probably would have resulted in different conclusions. However, the evidence obtained from the participants offers a broad understanding of how evaluation operates in both countries, the specific elements of the institutional framework that allow or hinder political influence, and, more importantly, how stakeholders from different spheres behave within this set of rules.

A second limitation is related to the scope of the research. In terms of evaluation and auditing, there is a vast range of policies, programmes, and organisations that have been assessed through different methods. Nonetheless, considering primarily the availability of resources to perform the fieldwork (time, access, funding, etc.) it was necessary to narrow the scope of analysis to two cases per country.

The development of the research implies developing strategies for coping with the limitations exposed in the pursuit of more robust findings. Regarding the limitation for generalisation of the case study, the incorporation of documentary analysis allows identifying structural elements of the evaluation policy in both countries, exceeding the boundaries of the case studies, and generating more data about the general state of evaluation. The identification of the institutional framework for evaluation/auditing, i.e. rules, procedures, and institutions, in both countries produced information about the hidden politics of evaluation applicable to programmes and organisations from other policy areas in these countries. In relation to the scope of the research, the selection of the case

studies was the result of a detailed process for identyfing those cases in the health sector with a wide range of evaluation outputs that could provide enough evidence for the analysis. The identification of interviewees also aimed to engage people from different sectors and backgrounds in obtaining different perspectives that could reflect the relationship between politics and evaluation beyond the particular reality of the cases analysed.

3.3.2 Challenges Faced During Fieldwork

In addition to the limitations associated with the methodological approach, the data-collection stage was challenging for a number of reasons. The most evident was that discussing about politics tended to create tension and reluctance in individuals when they feared that their responses might have a negative consequence for them, i.e. job loss. This is an element that is important to keep in mind because, as Desmond recognises, political sensitivity might affect the characteristics of the evidence provided by interviewees.[547]

Overcoming this challenge entailed a gradual process of relationship-building and the establishment of trust to make interviewees feel more confident for the discussion.[548] Based on this, it was a core element of the interaction with interviewees to continually transmit to them that their participation was highly valued and that their views and perceptions were respected. This resonated with the ethics underpinning the development of this research, especially in terms of how data was collected and the interaction with participants.[549] The sensitiveness of politics required being always particularly aware of the position of participants in disclosing information. Ethically, the key element was the transparency of the process and making explicit the purposes and expectations of their contribution. In line with the ethical principles identified by Burnham et al., the development of the research and, particularly, the interaction with interviewees were driven by the idea of minimising any

---

[547] Desmond, M. 2004. "Methodological challenges posed in studying an elite in the field." *Area* 36(3):262-69.
[548] Abma, T. A. 2006. "The Practice and Politics of Responsive Evaluation." *American Journal of Evaluation* 27(1):31-43.
[549] Miller, T., M. Birch, M. Mauthner, and J. Jessop (Eds.). 2012. *Ethics in Qualitative Research*. London, UK: SAGE.

potential harm derived from their participation, and respecting their privacy and confidentiality.[550] The core issue, as Seale states, 'is how the rights of participants (and researchers) are to be balanced against the potential benefits to society.'[551]

Another source of conflict was the identification and selection of possible interviewees. A clear limitation was that participation depended entirely on the will of each actor approached; some of them explicitly or implicitly (by not responding at all) declined the invitation for different reasons. Consequently, the research was based on the data collected only from those people who agreed to participate. This was a challenging situation. However, the networks established with some of the interviewees allowed the access to new potential participants to compensate for those who had declined the invitation. In some cases, it was also necessary to adjust to the limited time available for performing the interview because of the participant's busy agenda, a common issue with elite interviewees.[552] In terms of the documentary analysis, in both countries this was done, considering the information that was public and available. Some new documents were not considered because of the time framework of the research.[553] Table 3.3 refers to the deepened discussion about the challenges faced during the research process and the strategies applied to overcome these situations.

---

[550] Burnham, P., K. Gilland Lutz, W. Grant, and Z. Layton-Henry. 2008. *Research Methods in Politics.* Basingstoke, UK: Palgrave Macmillan.

[551] Seale, C. 2012. *Researching Society and Culture*. London, UK: SAGE. P. 73.

[552] Aberbach, J., and B. Rockman. 2002. "Conducting and Coding Elite Interviews." *Political Science & Politics* 35(4):673-76.

[553] For example a study about the CQC in progress.

Table 3.3 Main Challenges Faced During the Research Process

Challenge	Description	Strategy
Disperse information	During the documentary revision, the collection of the different evaluation/audit reports for each one of the case studies was complicated because this information is not concentrated into a single source.	Main sources of information were identified to make the collection and classification of data easier according to each case study. This allowed prioritizing those documents that were more relevant for the research.
Identification of potential interviewees	It was a difficult process, some of these actors who participated into the assessment of the programmes analysed have left their previous positions and they have new positions at present. Consequently, it was important to access to the people that were directly involved, as the focus of this research was the evaluation/audit processes of those programmes.	Some of the interviewees provided important information for contacting former officials or other people that could be relevant to interview.
Assymetries in the interviewing process	In the UK case, the process for inviting potential stakeholders to participate in the research was more time-consuming, some interviewees responded later than expected and this provoked delays in the process. In the case of the Mexican participants it was a more expedite process to establish contact; however, in the case of government officials, the end of the presidential term clashed with the scheduled interview sessions.	In some cases, it was necessary to modify the initial fieldwork programme. Particularly in the case of the UK, the supervisory team provided support for establishing contact with key potential interiewees.
Acces to high-level interviewees	In both countries, interviews performed to high-level stakeholders (e.g. politicians, government officials) were difficult to obtain and subject to the time-constrains imposed by their agenda.	The interviewing process was adapted, as possible, to these constraints to make interviewees aware of the relevance of their participation for this project.
Resistance to discuss 'the politics of evaluation'	In the Mexican case, some of the interviewees showed some resistance to opennly discuss the hidden politics of evaluation, they seemed careful not to disclose information that might be considered risky for them. In other cases, some participants seemed to be very interested in providing the 'institutional' view about the topic.[554] Some interviewees in the UK were at first reluctant to participate as it was perceived that as politics was the core issue to discuss, their views might be used for non-academic purposes.	Most of these resistance issues where overcome by establishing a relationship based on trust and confidence with interviewees, for example, by giving them more information about the nature of the research, the key issues to discuss and the treatment that will be given to the data they provide.

---

[554] See the concept of social desirability effect explained by Bryman. Bryman, A. 2012. *Social Research Methods*. Oxford, UK: Oxford University Press. ibid.

3.3.3 Other Considerations

The limitations derived from the methodological approach and the challenges faced during fieldwork are elements to contemplate when looking at the findings and conclusions presented in this thesis. Some other considerations are presented below:

- The qualitative approach is considered more susceptible to subjectivity. Therefore, it is not possible to completely abstract the professional bias of the researcher. Making explicit the research process undertaken reduces the potential negative influence of such a bias.
- Undertaking fieldwork in two different contexts might imply some asymmetry in the collection of data. Nevertheless it was a priority to carry out a very similar research process in both countries.
- Recommendations formulated for each case study are made on the basis of the evidence collected and the analysis undertaken. Its generalisation, as discussed in this chapter, is not the purpose of the research. Instead, the identification of potential new areas of knowledge for the evaluation studies field aligns best with this method.

The explanation of the process of this research is useful for having a broader understanding of how the data was collected, systematised, and interpreted. This chapter has also made explicit the limitations of the research and the methodological challenges that might have affected the outputs of this thesis. This information offers elements for identifying the scope of this research and the conditions underpinning it. The methodological framework detailed here serves as a basis for the description of the case studies selected.

The next two chapters present a careful description of the institutional framework that operates in the UK and in Mexico for evaluation and auditing. The case studies for the analysis of the hidden politics of evaluation are also addressed in terms of its functioning and of the evaluation processes that have been undertaken. These chapters aim to identify the set of rules existent in both countries that shape and constrain the interaction of the stakeholders of evaluation.

PART II
Empirical Case Study Research

Chapter 4. Case 1: Policy Evaluation in the United Kingdom

The theoretical grounds of this thesis (see Chapter 2) yield important elements for the comprehension of the political nature of evaluation. However, empirical evidence about this topic is still at a developing stage. There is an important gap in the knowledge in relation to how the political nature of evaluation manifests. This chapter describes how evaluation and auditing operate in the UK, to grasp how its relation to politics manifests in practice. The core argument presented here is that evaluation and auditing in the UK are systems that operate simultaneously but separately. The rules directed to this purpose give stakeholders an important amount of discretion to act affecting the overall effect of evaluation in governance. Consequently, the contribution of these instruments to policymaking is unclear and demands more careful examination of these processes. This description is mostly based on primary documents that allow identifying the rules and procedures for evaluation and auditing. The empirical evidence obtained through interviews is discussed later in the analytical chapter.

The chapter is divided into three parts. The first section sets out the rules and procedures that guide the performance of evaluation and auditing in the UK. The second section focuses on the first case study selected in the UK: the Cancer Strategy. This government initiative is taken as an element of analysis for the hidden politics of evaluation considering that it has been systematically evaluated and audited and it represents a central policy theme in the public agenda. The third section is dedicated to the second case study: the CQC. The analysis of this healthcare regulator is relevant, as it has been the result of a drive in the government for more efficiency. It also provides important knowledge about the outcomes of evaluation in relation to organisational performance. The core finding of this chapter is that policy evaluation in the UK is developed within an institutional framework that provides government departments with an important level of discretion in the decisions about *what*, *when* and *how* to evaluate. It is also characterised by a strong audit system led

by the National Audit Office (NAO) in which political strength is supported by the Select Committees of the House of Commons. These elements of the UK institutional framework contrast with the Mexican case where these activities are framed in a more rigid system of rules and procedures (see Chapter 5 below).

4.1 Audit and Evaluation in the UK

The UK has a long tradition in terms of monitoring and evaluation. Across time, different approaches have been implemented to develop accurate systems for evaluating policy. As Talbot observes:

> The UK has been successful in developing a fairly comprehensive performance and evaluation measurement, monitoring, and management system which, over time, has become increasingly "outcome" focussed, although at various levels it still includes large elements of output, process and even input monitoring.[555]

The reform of the public sector gave evaluation a more prominent role, particularly with the rise of the evidence-based policymaking approach.[556] At the beginning of the New Labour Government in 1997, this model became a recurrent element of the political discourse.[557] As Wells highlights, this fostered important changes in the public sector. On the one hand, evaluation gained significance as a systematic practice within central government and local agencies. On the other hand, evidence became a more frequent input for governmental tasks, along with VFM studies and outcome assessments.[558] The adoption of this model in the New Labour government can be seen as a point of departure for understanding the current institutional framework for evaluation and auditing in the UK.[559] The instruments generated for policy analysis aimed to obtain information about the effectiveness of programmes and

---

[555] Talbot, C. 2010a. "Performance in Government. The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom." Washington D.C., US: The World Bank. P. ix.

[556] European Commission. 2015. "Evidence-based Policymaking."

[557] See Cabinet Office. 1999a. "Modernising Government." London, UK.

[558] Wells, P. 2007. "New Labour and evidence-based policymaking: 1997-2007." *People, Place & Policy Online* 1(1):22-29.

[559] See: Solesbury, W. 2001. "Evidence Based Policy: Whence it Came and Where it's Going." ESRC Centre for Evidence Based Policy and Practice.

organisations.[560]

The growth of the audit field also influenced how policy is appraised and how data feeds back into policymaking. The 'audit explosion' described by Power made evident that a new dynamic between audited bodies and auditors was taking place,[561] similar to the relationship between evaluators and evaluands. The shift towards a more rational public management and the use of tools like monitoring and evaluation created new conditions in the political arena.[562] The hidden politics of evaluation matter because of the analysis of those relationships and their consequences. The political nature of evaluation concerns the way in which the stakeholders involved (practitioners, politicians, evaluators, etc.) make use of their resources to pursue their interests. As the bond between evaluation and politics can produce both positive and negative outcomes, it is important to bear in mind how these elements interact.

From an overall perspective, it is possible to identify in the UK two major systems: internal evaluation (performance monitoring and evaluation studies) and an external audit system (Table 4.1). These systems are underpinned by different values, pursue different purposes, and are based on different processes. The values behind evaluation and auditing will inevitably affect the formal and informal arrangements produced. This generates a series of reciprocal reactions in which these processes are affected by the stakeholders' values, while those values are also shaped by the conditions in which these

---

[560] The DfID is a clear example of a central department in which evaluation is a systematic and internalised practice. See: Department for International Development. 2014. "Evaluations completed: April 2014 to March 2015." London, UK.

[561] Power, M. 1997. *The Audit Society: Rituals of Verification*. Oxford, UK: Oxford University Press.

[562] In addition to the 'audit explosion' identified by Power, the rise in the use of evaluation mechanisms in the public sector has been also studied by various scholars in the field. For example, Pollitt considers that in Europe the furore caused by policy evaluation 'is taking place not because executive politicians have suddenly changed their natures, but because their environments have shifted. In particular, the choices they face are more tightly resource constrained and they can no longer call on the same reserves of legitimacy and authority which were available to them two or three decades ago.' Pollitt, C. 1998. "Evaluation in Europe: Boom or Bubble?" *Evaluation* 4(2):214-24. P. 223.
The way in which both audit and evaluation have gained spaces in the public sector is underpinned not only by a managerial logic seeking more efficiency, but also by a political rationale that cannot be ignored. See: Datta, L. E. 2001. "Coming attractions." *The American Journal of Evaluation* 22(3):403-08, Mark, M. M. Ibid."Evaluation's future: furor, futile, or fertile?":457-79. , Pollitt, C. 1998. "Evaluation in Europe: Boom or Bubble?" *Evaluation* 4(2):214-24.

activities occur. The main element that this research aims to analyse is the way in which institutions affect the hidden politics of evaluation. Looking at how this variable engages in different contexts will detect those conditions that favour or limit the establishment of power relationships and the implications for the objectivity, independence, and utility of evaluation processes.

Table 4.1 Evaluation and Audit Processes in the UK

Internal Evaluation	External Auditing
• Performance monitoring activities through the use of indicators and the follow-up of organisational goals, overviewed by HM Treasury • No integrated evaluation policy, only guidelines for its development (the *Green Book* and the *Magenta Book*) • Evaluation studies carried out by government departments through internal units or external consultants	• Audit system led by the NAO • The NAO scrutinises public spending on behalf of Parliament • It performs Financial Audits and VFM studies, followed-up by the Public Accounts Committee (PAC)

The institutional architecture for evaluation in the UK will reveal the formal procedures that constrain its performance, particularly those areas of the process with political interests at stake, i.e. influencing the selection of methods, as well as the use given to evaluation findings for the legitimisation of public action.

4.1.1 The Institutional Framework for Policy Evaluation in the UK

The relationships that emerge from assessment processes are constrained by the limits and boundaries set by formal and informal rules. The institutional framework does not refer only to the legislation and organisations involved in evaluation and auditing, but also to the set of values (explicit or not) underpinning these activities (see Chapter 2, Section 2.2.3). The UK has a long tradition of initiatives for the assessment of 'what works.'[563] As stated, it is possible to identify two major processes for evaluation in the UK: internal evaluation (studies developed *within* departments or commissioned to external

---

[563] Talbot, C. 2010a. "Performance in Government. The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom." Washington D.C., US: The World Bank.

parties), and the *external* auditing performed by the NAO on behalf of Parliament. These activities are described in more detail in subsequent sections. However, it is worth providing an overview of the system operating in the UK. At the national level, evaluation research is carried out by government departments through internal units or external consultants. The Department of Health (DoH), for example, systematically produces evidence to support decision-making.[564] In terms of performance monitoring, central government departments continually report a set of indicators (input and impact) about the activities carried out for addressing governmental priorities.[565] The audit system is conducted by the NAO, an organisation that assesses central government departments in terms of efficiency, effectiveness, and economy and represents the pillar on which the accountability system is based.

Evaluation in this country is entrenched in the public sector and 'literally hundreds of evaluations are carried out.'[566] To illustrate, the DoH recently commissioned a study about its relationship with stakeholders and their perceptions of an external organisation.[567] This suggests that evaluation within central government involves stakeholders from different sectors to increase the robustness and legitimacy of evaluation.[568] The most distinctive feature of the UK case is that there are no explicit obligations for central government departments to evaluate their programmes. Instead, there are guidelines that departments might or might not follow.[569] According to the NAO, 'Departments vary in the extent to which they follow central and department-specific

---

[564] See: Advisory Committee on the Safety of Blood, Tissues and Organs. 2014. "Transplantation of Organs from Deceased Donors with Cancer or a History of Cancer." Department of Health.

[565] See for example: Department of Health. 2010. "Business Plan." London, UK.

[566] Schiavo-Campo, S. 2005. "Building Country Capacity for Monitoring and Evaluation in the Public Sector: Selected Lessons of International Experience." edited by The World Bank. Washington, D.C., US. P. 12.

[567] Department of Health. 2013a. "Department of Health 2013 Stakeholder Research." London, UK.

[568] See also: Department for Culture, Media and Sport. 2013. "An Independent Review of E-Lending in Public Libraries in England." London, UK, Department for International Development. 2010. "Joint Irish Aid and DFID's country programme evaluation Tanzania 2004/05 to 2009/10." London, UK.

[569] The guidelines for central government departments comprise multiple approaches and methods for evaluating public policy. See: HM Treasury. 2011c. "The Magenta Book. Guidance for evaluation." edited by HM Treasury. London, UK.

requirements on cost-effectiveness evaluation.'[570] Although guidelines for evaluation exist, these are not necessarily followed. Government departments have space for discretionary decision-making in evaluation, for example, in terms of the approaches applied and the stakeholders involved. This can be observed in the health sector. Organisations like the National Institute for Health and Care Excellence (NICE) and Public Health England (PHE) are systematically involved in evaluation by issuing guidelines, piloting initiatives, or developing frameworks.[571]

According to Henkel, the promotion of evaluation as a mechanism for strengthening policymaking in the UK is related to three major purposes: 'To control public expenditure, to change the culture of the public sector and to shift the boundaries and definition of public and private spheres of activity.'[572] Clarke agrees by stating:

> Establishing, assessing, and improving the performance of public services has become an increasing obsession of governments and institutions of evaluation have been adapted and invented to embody this obsession [...] Britain has been a leading force in developing this approach that simultaneously builds on and transforms well-established approaches to audit and inspection.[573]

This blooming of evaluation can be associated with the rise of the NPM. The pressures faced by the government to make more with less exposed the need for more efficient and effective policies. During the New Labour government (1997–2007), for example, the modernisation of the public sector was a central element of the political discourse.[574] This can be observed in the *Modernising Government* White Paper (1999):

> Our aim is to reallocate money to key priorities; to change policies so that

[570] National Audit Office. 2013a. "Evaluation in government." edited by National Audit Office. London, UK. P. 13.

[571] PHE, for instance, has developed standard evaluation frameworks for weight management, physical activity and dietary interventions. See: Public Health England. 2013. "Standard Evaluation Frameworks." London, UK.

[572]Henkel, M. 1991. *Government, Evaluation and Change*. London, UK: Jessica Kingsley Publishers. P. 9.

[573] Clarke, J. 2005. "Performing for the Public: Doubt, Desire, and the Evaluation of Public Services." in *The Values of Bureaucracy*, edited by P. Du Gay. Oxford, UK: Oxford University Press. P. 212.

[574] See also: Finlayson, A. 1998. "Tony Blair and the jargon of modernisation." *Soundings Autumn* (10).

> money is well spent; to ensure that departments work better together to improve services; to weed out unnecessary or wasteful spending.[575]

> Modernising government means [...] raising all standards until they match the best within and outside the public service, and continue improving. It means transforming government, so that it is organised around what the public wants and needs, rather than around the needs or convenience of institutions.[576]

The adoption of NPM reforms pushed into the government's modernisation agenda the need to measure the effectiveness of public policy by 'creating a culture in which public bodies are continually reflecting on the services they provide and how they may be improved.'[577] This became a general concern for the government, as there was an increasing necessity to be more parsimonious in the expenditure of public funds, along with a stronger social demand for transparency and accountability. The *Open Data* White Paper, for example, highlighted the relevance of a more transparent policymaking process: 'We are proactively releasing much of the underlying, objective data used by the public sector to arrive at policy recommendations and in supporting policy decisions.'[578] The promotion of evaluation and audit mechanisms has been an important element both within the Executive and the Legislative.

Instruments for performance monitoring and evaluation grew both in the local and national spheres.[579] What can be labelled as 'evaluation policy' in the UK is the co-existence of a performance monitoring system (implemented by HM Treasury), the use of evaluation research by government departments, and an audit system driven by the NAO. This is translated into the implementation of parallel processes that generate multiple outputs. As Dorey explains:

> The increased recourse to evaluation both reflects and reinforces a belief

---

[575] Chancellor of the Exchequer. 1998. "Modern Public Services for Britain: Investing in Reform. Comprehensive Spending Review: New Public Spending Plans 1999-2002."

[576] Prime Minister and the Minister for the Cabinet Office. 1999. "Modernising Government."

[577] McAnulla, S. 2006. *British Politics: A Critical Introduction*. London, UK: Continuum. P. 151.

[578] HM Government. 2012. "Open Data White Paper. Unleashing the potential.", edited by Cabinet Office. London, UK. P. 31.

[579] Bjørnholt, B., and F. Larsen. 2014. "The politics of performance measurement: 'Evaluation use as mediator for politics'." *Evaluation* 20(4):400-11, Haveri, A. 2008. "Evaluation of Change in Local Governance: The Rhetorical Wall and the Politics of Images." Ibid.14(2):141-55, Martin, S., J. Downe, C. Grace, and S. Nutley. 2010. "Validity, Utilization and Evidence-Based Policy: The Development and Impact of Performance Improvement Regimes in Local Public Services." Ibid.16(1):31-42.

among many policy makers that it is possible – and necessary, for reasons of efficiency and effectiveness – to develop a more rational and scientific approach to developing policies, based both on measuring statistical data and conducting experiments via 'pilot schemes'[580] and consequently determining 'what works.'[581]

The UK institutional framework is relevant in terms of how it affects the political relationships that are established as a consequence of evaluation. The description of each one of these activities elucidates those elements that are crucial for the analysis of the hidden politics of evaluation.

4.1.2 Internal Evaluation

As discussed, internal evaluation comprises two main processes: performance monitoring and policy evaluation. Performance monitoring is a much more institutionalised practice, promoted—through a range of instruments—by different parties in the government.[582] It has comprised a vast range of tools that has evolved over time, but that have remained consistently at the core of central government.[583]

The main reason for the predominance of this approach is the construction of a linkage between performance and the allocation of public funds.[584] This reveals an explicit focus on the promotion of efficiency and effectiveness in the delivery of public services.[585] During the New Labour government, measuring

---

[580] Recently it has been debated in the UK the pertinence of carrying out policy pilots in contrast to other evaluation mechanisms like randomised –controlled trials (RCT). This comes from the idea that a pilot can effectively provide information about the feasibility of implementing a policy, but are not able to provide evidence about the potential harms of benefits that a policy can have because contextual elements change. Goldacre, B. 2013. "Building evidence into education." London, UK.

[581] Dorey, P. 2014. *Policy Making in Britain: An Introduction*. London, UK: SAGE. P. 280.

[582] A chronology of performance management systems in the UK developed by Bourn (2007) is shown in *Appendix D*.

[583] For a detailed description of the evolution of the performance monitoring system in the UK see Talbot, C. 2010a. "Performance in Government. The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom." Washington D.C., US: The World Bank.

[584] Regarding this issue, the work of Curristine 'provides an overview of the development and use of performance measures and evaluations in the budget process across OECD countries.' This study highlights, for example, the case of the Research Assessment Exercise in the UK, where performance information was linked to budgeting processes. This instrument was substituted by the Research Excellence Framework in 2014. See: Curristine, T. 2005. "Performance Information in the Budget Process: Results of the OECD 2005 Questionnaire." *OECD Journal on Budgeting* 5(2):87-131. P. 87.

[585] The distortions produced by the development of performance systems have been discussed by Clarke, J. 2008. "Performance Paradoxes: The Politics of Evaluation in Public Services." Pp.

performance was a central issue and HM Treasury implemented the use of *Public Service Agreements* (hereafter, PSA).[586] These were essentially 'contracts' between HM Treasury and departments based on the accomplishment of performance targets and "if not met the Treasury will 'claw back' funds they would otherwise release to departments."[587] In the 1998 Comprehensive Spending Review, the Prime Minister (PM) emphasised the role of these instruments in the government's modernisation agenda:

> It reflects detailed public service agreements between departments and the Treasury about how these overall totals will be spent, with clear objectives and output and efficiency targets that departments have agreed to meet.[588]

The establishment of PSA reveals an explicit bond between public spending and efficiency. PSAs were founded on a set of principles to promote a more detailed review of public expenditure and incorporated 'high-level objectives and targets which commit the department to achieving a specific level of success.'[589] Also, as James highlights, PSAs were 'a tool for negotiating changes with individual departments, [...] to improve the co-ordination of priority setting where policy or delivery issues cut across departmental boundaries.'[590] Matthews describes the characteristics of this system:

- Link between expenditure and performance was made explicit; resources were allocated in accordance with the Government's broad aims, translated into departmental objectives, against which PSA targets were set
- Shift from annual budgeting to three-year Department Expenditure Limits
- Longer-term planning
- Targets were intended to be SMART (specific, measurable, achievable, relevant, and timed)

120-34 in *Public Services Inspection in the UK. Research Highlight in Social Work*, edited by H. Davis and S. Martin. London, UK: Jessica Kingsley Publishers.

[586] PSAs can be considered a substitute mechanism of the *Output and Performance Analysis (OPA)* established during the Conservative Government during the 1990s. For further reference see: Matthews, F. 2013. *Complexity, Fragmentation, and Uncertainty. Government Capacity in an Evolving State.* Oxford, UK: Oxford University Press. Talbot, C. 2010a. "Performance in Government. The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom." Washington D.C., US: The World Bank.

[587] McAnulla, S. 2006. *British Politics: A Critical Introduction*. London, UK: Continuum. P. 152.

[588] Chancellor of the Exchequer. 1998. "Modern Public Services for Britain: Investing in Reform. Comprehensive Spending Review: New Public Spending Plans 1999-2002."

[589] Johnson, C., and C. Talbot. 2007. "The UK Parliament and performance: challenging or challenged?" *International Review of Administrative Sciences* 73(1):113-31. P. 116.

[590] James, O. 2004. "The UK Core Executive's Use of Public Service Agreements as a Tool of Governance." *Public Administration* 82(2):397-419. P. 400.

- Encouragement of cross-departmental work
- Greater degree of autonomy for front-line service providers
- Publicity of performance information.[591]

PSAs also aimed to incentivise departments to achieve their organisational goals during a specific period of time and to give them more flexibility in the administration of their budget …

> … to improve the coordination of priority setting where policy or delivery issues cut across departmental boundaries […] it was designed to bring all of central government under a system-wide performance regime to reduce fragmentation.[592]

However, the effectiveness of the PSAs was not exempted from criticism. The Institute for Government explains that after the first New Labour term …

> … Blair was frustrated with the lack of progress made during his first term and wanted to use his second term to push through radical public service reform […] the PSA agenda became significantly more ambitious, but more focused.[593]

The second New Labour term brought important changes to this scheme. The number of targets was reduced from 600 (in 1998) to 160 (in 2000)[594] and other modifications were made (mostly related to the specification of indicators).[595] In the British public sector, the PSA system …

> … dominated the government's approach to monitoring public sector performance for 12 years and played a key role in evaluating the progress of public sector reform, but it was criticised for being over-centralist and bureaucratic despite being refined and simplified over the years while some targets proved counter-productive.[596]

The identification of goals and the construction of indicators promoted by the PSA system fit better under the label of monitoring rather than under evaluation (see Chapter 2, Section 2.1.3). The main function of instruments of this nature is having a better control over the tasks given to each department and the goals

---

[591] Matthews, F. 2013. *Complexity, Fragmentation, and Uncertainty. Government Capacity in an Evolving State.* Oxford, UK: Oxford University Press. Pp. 84-86.

[592] James, O. 2004. "The UK Core Executive's Use of Public Service Agreements as a Tool of Governance." *Public Administration* 82(2):397-419. P. 400.

[593] Panchamia, N., and P. Thomas. 2014. "Public Service Agreements and the Prime Minister's Delivery Unit." Institute for Government. P. 3-4.

[594] Gay, O. 2005. "Public Service Agreements." Parliament and Constitution Centre. P. 2.

[595] Department of Health. 2006a. "Autumn Performance Report." London, UK.

[596] Burton, M. 2013. "The Politics of Public Sector Reform." Hampshire, UK: Palgrave Macmillan. P. 228.

achieved. The benefits produced by this system were recognised even beyond the government's sphere. The Select Committee of Public Administration, for example, acknowledged that PSAs allowed …

> … targeting shifts attention from the classic Treasury concerns of inputs (money and personnel) to outputs and outcomes.[597]

Efficiency and the public sector reform were still priority topics in the agenda of the Coalition Government (2010–15), as it can be observed in the statement made by the PM, David Cameron, during the Confederation of British Industry's Conference:

> We had to cut the cost of government, we had to get the deficit down and we've got to live within our means. Now, we've made good progress but there's a huge amount left to do […] we've managed to find £14.5 billion of efficiencies within government – that's just about doing things better.[598]

Despite the commitment to these values, the PSA system was abolished and a new performance-measurement system was implemented by the Coalition Government: the *Business Plans*. These "placed considerable emphasis on 'milestones' against which to determine departments' success in delivering commitments, with reporting structures leading directly to Downing Street."[599] These documents set out …

> … coalition priorities, key policy and implementation actions for the remainder of this Parliament; expenditure for each remaining year of this Spending Review period, and other departmental data, including an organogram, input and impact indicators.[600]

Business Plans enable monitoring the level of accomplishment made by departments and include key information about the connection between the Coalition priorities and departmental activities:

> Vision and priorities to 2014-15;
> Structural reform plan, including actions and deadlines for implementing reforms over the next two years; and

---

[597] Public Administration Select Committee. 2003. "On Target? Government By Measurement." London, UK. P. 9.

[598] Cabinet Office. 2014a. "CBI Annual Conference 2014: Prime Minister's address.".

[599] Painter, C. 2012. "The UK Coalition government: Constructing public service reform narratives." *Public Policy and Administration* 20(1):3-20. P. 9.

[600] British Prime Minister's Office. 2010. "Business Plans: track progress in implementing our policies." London, UK.

Contribution to transparency, including the key indicators against which it will publish data to show the cost and impact of public services and departmental activities.[601]

The establishment of indicators has an important political implication to highlight. Quantifiable measurements are an asset for justifying policy decisions because the conceptualisation of these instruments is underpinned by values like objectivity and political neutrality. As The World Bank recognises, 'With indicators, monitoring and evaluation is more compelling because it is objective, not based on personal judgments or pure description.'[602]

An interesting example of this can be found in the health sector. The Coalition Government established six priorities for the DoH. One of these priorities was 'to enable better health and wellbeing for all.' There are actions and indicators associated with this priority that are systematically monitored by the Department and informed to HM Treasury (Box 4.1).[603] This disaggregation of political priorities into concrete actions and quantifiable indicators is intended to generate more elements to control performance and to report progress.

Box 4.1 Example of the Monitoring of Coalition Priorities in the Health Sector

> **Coalition Priority**
> To enable better health and wellbeing for all.
> **Action**
> Reduce preventable early death.
> **Indicator**
> Mortality rate from causes considered preventable

Source: Department of Health, 2010. *Business Plan.*

Business plans are also a core element of the planning process. The development of the DoH's Corporate Plan is closely linked to the targets and indicators of their Business Plan. Following the example presented in Box 4.1, the Corporate Plan disaggregates the governmental priorities and establishes specific actions to implement during a period of time. For the priority 'to enable better health and wellbeing for all', this document recognises the need for

---

[601] HM Treasury. 2010. "Spending Review 2010." edited by HM Treasury. London, UK. P. 9.
[602] Mosse, R., and L. E. Sontheimer. 1996. "Performance Monitoring Indicators Handbook." Washington, D.C., US: The World Bank. P. 4.
[603] Regarding this issue, in 2014 The Institute for Government analysed how departments measure performance and the use of Business Plans. See: Freeguard, G., P. Bouchal, R. Munro, C. Nimmo, and J. McCrae. 2014. "Whitehall Monitor 2014. A data-driven analysis of the size, shape and performance of Whitehall." London, UK: Institute for Government.

'preventing people from dying prematurely by improving mortality rates for the big killer diseases, to be the best in Europe through improving prevention, diagnosis and treatment and reducing health inequalities.'[604] This shows that the use of performance information is embedded in the daily operation of departments and is an input for reporting progress.[605]

This focus on performance monitoring has led to the development of other instruments of control. For instance, the 2013/14 NHS's Outcome Framework[606] 'contains a number of indicators selected to provide a balanced coverage of NHS activity.'[607] This allows the NHS to adapt the planning and monitoring of its activities to those priority-policy themes in the agenda. The NHS's framework, for instance, incorporated new indicators for cancer to improve the monitoring of the services provided, e.g. five-year survival from all cancers. This suggests that planning processes are flexible to adapt to political changes in the priorities.

In addition, departments issue a Quarterly Data Summary (QDS) that provides an update on how the budget is spent. It is mostly a transparency action that gives taxpayers a synthetic view of their financial performance: 'To make more of the management information currently held by government available to members of the public on a regular basis.'[608] Unlike other tools like evaluation and auditing, performance monitoring lacks the judgement component about policy. Nonetheless, monitoring data can represent an input for evaluation and auditing in terms of the interpretation that is made from this evidence.

Policy evaluation has had a more fragmented development in the UK. Even though efforts in this area have been made, especially in health and education

---

[604] Department of Health. 2014a. "DH Corporate Plan 2014- 2015." edited by Department of Health. London, UK. P. 11.

[605] See for example: —. 2014b. "Mid-year report 2014-15." London, UK.

[606] The purpose of this framework is to 'provide a national overview of how well the NHS is performing; be the main way, along with the NHS Mandate, in which NHS England is held to account for improvements in health outcomes; improve quality throughout the NHS by encouraging a change in culture and behaviour focused on health outcomes not process.' —. 2013b. "Health and social care outcomes frameworks." London, UK.

[607] National Health Service. 2012. "The NHS Outcomes Framework 2013/14." edited by National Health Service. London, UK. P. 4.

[608] Department of Health. 2012a. "DH business plan quarterly data summary: January 2012." London, UK.

sectors, the rise of evaluation has been a more recent process.[609] According to Pollitt, in the UK …

> … attempts have certainly been made to install evaluation at the core of central government […] yet, despite initial enthusiasms, the institutional homes built for policy evaluators have, one by one, been demolished by the huff and the puff of political forces and intellectual fashions.[610]

With the arrival of the New Labour government, the interest for evidence-based policymaking rose as a strategy to give more rationality to policy decisions.[611] The enthusiasm for this model was even considered by some scholars as 'somewhat of a mantra for governments.'[612] Under the administration of the PM, Tony Blair, the *Modernising Government* White Paper claimed for a 'better use of evidence and research in policymaking and better focus on policies that will deliver long-term goals.'[613] Its use in the public sector 'was intended to signify the entry of a government with a modernising mandate, committed to replacing ideologically-driven politics with rational decision making.'[614] This gave evaluation a more salient role, as 'having to specify precise outcomes products and milestones early on helps policy makers to build evaluation into the policy making process from the outset.'[615] However, evaluation exceeds the framework of the evidence-based policymaking model because the use of its outputs is not constrained to support decision-making. These might also contribute to broader aspects such as transparency and learning.[616]

---

[609] Talbot, C. 2010a. "Performance in Government. The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom." Washington D.C., US: The World Bank.

[610] Pollitt, C. 1993. "Occasional Excursions: A Brief History of Policy Evaluation in the UK." *Parliamentary Affairs* 46(3):353-62. P. 354.

[611] For further reference see: Bogenschneider, K., and T. J. Corbett. 2010. *Evidence-Based Policymaking: Insights from Policy-Minded Researchers and Research-Minded Policymakers*. New York, US: Routledge, Guenther, J., E. Williams, and A. Arnott. 2010. "The politics of evaluation: evidence-based policy or policy-based evidence?" in *NARU Public Seminar Series*. Darwin, Australia, McDavid, J. C., and I. Huse. 2012. "Legislator Uses of Public Performance Reports." *American Journal of Evaluation* 33(1):7-25, Shaw, I., J. C. Greene, and M. M. Mark. 2006. *Handbook of Evaluation: Policies, Programs and Practices*. London, UK: SAGE.

[612] Vaessen, J., and F. L. Leeuw. 2010. *Mind the gap: Perspectives on Policy Evaluation and the Social Sciences*. New Jersey, US: Transaction Publishers. P. 136.

[613] Cabinet Office. 1999a. "Modernising Government." London, UK.

[614] Sutcliffe, S., and J. Court. 2005. "Evidence-Based Policymaking: What is it? How does it work? What relevance for developing countries?", edited by Overseas Development Institute. P. 1.

[615] Cabinet Office. 1999c. "Professional Policy Making for the Twenty First Century." P. 18.

[616] Segone, M. (Ed.). 2008. *Bridging the gap. The role of monitoring and evaluation in evidence-based policy making*: UNICEF.

The promotion given to evaluation by the New Labour government represented, at least discursively, a linkage between evidence and accountability and a high-priority issue in the government's agenda.[617] During this period, it is possible to find recurrent allusions to evidence and evaluation, both in terms of its potential benefit for policymaking, as in terms of the need to develop more capacities in this area:

> -Every new initiative needs a built-in capacity to learn from monitoring and evaluation.[618]
> -A more rigorous approach to appraisal and evaluation is needed so that information requirements are sufficient to enable Government to appraise policy options, monitor progress and evaluate results.[619]

This focus on evidence made the connection between social-science research and policy more explicit. According to Wells, this fostered the development of large-scale evaluations to social policies.[620] The Sure Start initiative,[621] for example, was profusely evaluated using multiple methodologies and generated important evidence about the effectiveness of the programme.[622] The design and implementation of evaluation tools in different policy areas became also more frequent.[623] This interest on evidence can be seen as a way for policymakers to set themselves apart from the politics of the public sector by relying on objective and so-called 'neutral information.'[624] Nonetheless, as will

---

[617] Monaghan, M. 2011. *Evidence versus Politics: Exploiting Research in UK Drug Policy Making?* Bristol, UK: Bristol Policy..

[618] Mulgan, G., and A. Lee. 2001. "Better Policy Delivery and Design: A Discussion Paper." edited by Performance and Innovation Unit. London, UK.P. 18.

[619] Cabinet Office. 2000. "Wiring it up: Whitehall's Management of Cross Cutting Politics and Services." London, UK. P. 38.

[620] Wells, P. 2007. "New Labour and evidence-based policymaking: 1997-2007." *People, Place & Policy Online* 1(1):22-29.

[621] These programmes 'were aimed at families with children up to the age of 4 living in disadvantaged areas [...] the aim was to bring together early education, childcare, health services and family support to promote the physical, intellectual and social development of babies and children.' Source: Department for Education. 2011. "National evaluation of Sure Start local programmes: an economic perspective." London, UK.

[622] See: Birbeck University of London. 2012. "National Evaluation of Sure Start." London, UK.

[623] Sutcliffe and Court review some of the specific tools used by the UK government to obtain evidence, including *The Green Book* and *The Magenta Book* described later in this chapter. Sutcliffe, S., and J. Court. 2005. "Evidence-Based Policymaking: What is it? How does it work? What relevance for developing countries?", edited by Overseas Development Institute.

[624] This argument can be related to the idea of *depoliticisation* discussed in Chapter 2. From this perspective, evaluation and audit can be seen as mechanisms to shift the political values associated to public policy through the identification of evidence obtained through so-called objective and unbiased methods. As it will be discussed in Chapter 6, the political values embedded in evaluation inevitably manifest during different stages of the process. Therefore,

be discussed later, this is not exempt from political influence. The flourishing of evaluation during the New Labour's government promoted a new dynamic between the public sector and the market of knowledge, which, according to Sullivan, did not have an important impact on policymaking at the time:

> Despite the considerable investment in policy evaluation, 'evidence-based policymaking' did not come to define the policy process but rather generated a disconnection between 'evidence' and 'argument' that diminished the role of the evaluator in the policy process.[625]

As stated, the evidence-based approach was a core element of the political discourse, but in practice, the linkage between social-science research and public policy was not so perceivable. Parsons observes:

> The speech turned out, therefore, to be less of a new chapter in the history of public policy, so much as 'deja vu all over again'[...] If 'evidence' is to really to have more of an influence on policy-making then this involves tackling the thorny Lasswellian question of 'who gets what, when and how'. Whose evidence gets what influence, when and how?[626]

Consequently, the impact of the evidence-based approach promoted by New Labour is debatable. On the one hand, its adoption contributed to the development of evaluation capacities in central government.[627] On the other hand, its effect on policymaking is unclear and it requires assessment in terms of its utility for supporting and legitimising this process. According to Sutcliffe and Court, the UK 'needs to be increased communication and interaction between the research and policy worlds in order to strengthen the integration of policy and evidence.'[628] This makes evident the complexity of evaluation, not only in terms of establishing guidelines for its performance, but also to generate incentives to promote its utilisation.

---

depoliticisation can be seen as a process in which politics are removed from a highly technical area like methodology, to provide an example, and are moved to other areas like the definition of the scope or the dissemination of results. See: Flinders, M., and J. Buller. 2006. "Depoliticisation: Principles, Tactics and Tools." *British Politics* 1(3):293-318.

[625] Sullivan, H. 2011. "'Truth' junkies: using evaluation in UK public policy." *Policy & Politics* 39(4):499-512.P. 510.

[626] Parsons, W. 2002. "From Muddling Through to Muddling Up - Evidence Based Policy Making and the Modernisation of British Government." *Public Policy and Administration* 17(3):43-60. P. 57.

[627] Davies, P. 2012. "The State of Evidence-Based Policy Evaluation and its Role in Policy Formation." *National Institute Economic Review* 219(1):R41-R52.

[628] Sutcliffe, S., and J. Court. 2005. "Evidence-Based Policymaking: What is it? How does it work? What relevance for developing countries?", edited by Overseas Development Institute. P. 14.

In the Coalition Government (2010–15) there was no integrated evaluation policy or regulation that enforced evaluation; instead, certain guidelines were made available to departments.[629] According to an evaluator (UK-14):

> In theory, any discretionary spending that the Treasury enable a department to engage in is required to do some kind of evaluation, but there are no sort of binding conditions on precisely how that evaluation is undertaken.

Although there are guidelines for the performance of evaluation, each department has its own internal procedures for this activity—except for the public bidding of evaluation contracts.[630] Despite evaluation not being mandatory, departments have systematically used social science research in different ways.[631] According to Morris, evaluation in the UK has been mostly used for the following purposes:

- Testing of new policy ideas – pilots and demonstrations (limited to specific areas/time periods)
- Identifying whether existing interventions need to be reformed or abandoned
- Improvement of implementation and management of interventions
- Development of new interventions – action research
- Pragmatic – politicians want evidence that their programmes/interventions work as a political defence.[632]

The use of pilots was recurrent within the Coalition Government for the identification of the economic impact of a new policy.[633] These aim to reduce the possible negative effects produced by executing a policy that might not work. It 'will be a prelude to the policy being implemented more widely or nationally,

---

[629] The Government Social Research Service (GSR) has issued guidance for evaluation about multiple topics (methodology, ethics, and policy analysis). See: Government Social Research Service. "Professional Guidance." London, UK.

[630] Public organisations must follow the procedures and directions established in the *Managing Public Money* book issued by HM Treasury. This document 'sets out the main principles for dealing with resources in UK public sector organisations.' This includes the guidelines for public organisations to follow when they need to hire the services of external consultants. HM Treasury. 2013. "Managing Public Money." edited by HM Treasury. London, UK. P. 8.

[631] For example, the *Behavioural Insights Toolkit* study performed by the Social Research and Evaluation division in the Department for Transport and the *Climate Change and Transport Choices* study commissioned by the Department of Transport to TNS BMRB, an external consultant.

[632] Morris, S. 2005. "Evaluating Public Policy in the UK: History, Politics & Practice." P. 7.

[633] During the last years, an important number of evaluation studies for testing new policies have been performed by different departments. For example, last year, the Ministry of Justice commissioned to a group of researchers from *The University of Kent* the evaluation of the mandatory polygraph pilot. The DoH also commissioned a study to the NatCen Social Research and the University of York for the Drug System Change Pilots Evaluation.

once its effectiveness has been evaluated and any problems have been addressed.'[634] It can also be seen as a way to minimise the potential political costs of failure. Other types of recurrent evaluation in this country are those focused on implementation, useful for detecting those potential areas of improvement.[635]

Another relevant issue during the Coalition Government (2010–15) was the uneven evaluation capabilities among departments. This can be explained by the lack of a structured evaluation policy that standardises its performance. For example, there is no integrated information about the number, nature, and cost of evaluations.[636] There are also important variances in the number of studies undertaken and most departments only make evaluation reports public without providing any further information.[637] Despite this, some departments present a more developed evaluation strategy. The Department for International Development (DfID), for instance, has an explicit annual evaluation policy that sets out the ethical principles for research and evaluation. It also delivers annually a report about the evaluations undertaken (approaches, methods, and resources).

Guidelines for evaluation in central government are concentrated in two instruments issued by HM Treasury. The first one is *The Green Book,* published during the New Labour government and revised by the Coalition Government.[638] This document established that 'all new policies, programmes and projects, whether revenue, capital or regulatory, should be subject to comprehensive but proportionate assessment, wherever it is practicable, so as best to promote the

---

[634] Dorey, P. 2014. *Policy Making in Britain: An Introduction*. London, UK: SAGE. P. 279.

[635] Examples of this kind of evaluation are the Evaluation of the SEND pathfinder programme (Process and Implementation Research Report) and the evaluation of the 'Travelling to School Initiative' Programme, both commissioned by the Department of Education.

[636] One of the criticisms made by the NAO in its report *Evaluation in government* published was precisely the lack of guidance for departments 'on how to implement the principle of proportionality in what and how to evaluate'. National Audit Office. 2013a. "Evaluation in government." edited by National Audit Office. London, UK. P. 13. *Appendix F* summarises the key findings of this study.

[637] For example during 2014, the Department for Work and Pensions performed more than 20 studies, while the Department for Transport less than five.

[638] The current edition of *The Green Book* was issued in 2003, with some minor changes done in 2011. According to the HM Treasury's website, *The Green Book* is going through a process of revision.

public interest.'[639] Some of the most relevant topics addressed in this handbook are:

- Appraisal and evaluation as part of the policy cycle understood within central government as ROAMEF—an acronym for: rationale, objectives, appraisal, monitoring, evaluation, and feedback (Figure 4.1).
- It distinguishes between these two concepts by describing appraisal as an activity prior to the implementation of a policy, and evaluation as an *ex-post* judgement.
- It provides methodological tools for assessing the pertinence of the government's intervention.
- Appraisal focuses on the performance of *cost-benefit analyses*; *The Green Book* provides a solid body of methodological tools for this purpose.
- Evaluation is seen as a mechanism to compare the results obtained by a policy, programme, or project in relation to what it was expected to achieve.

Figure 4.1 Rationale, Objectives, Appraisal, Evaluation and Feedback (ROAMEF) Cycle
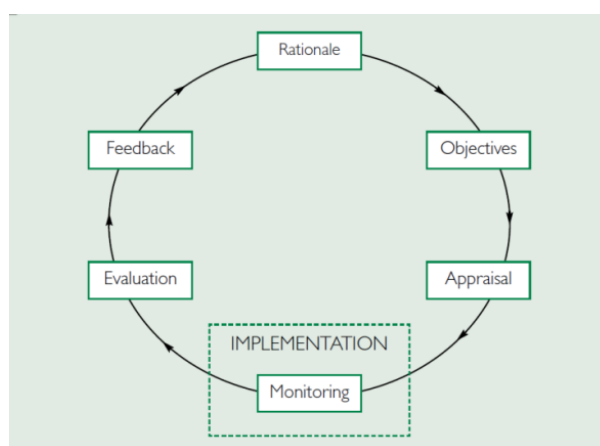
Source: HM Treasury, 2011. *The Green Book.* P. 3

The Green Book distinguishes between appraisal and evaluation and this has derived into two different processes (Figures 4.2 and 4.3).

---

[639] HM Treasury. 2011a. "The Green Book. Appraisal and Evaluation in Central Government." London, UK.

Figure 4.2 *The Green Book* Process for the Appraisal of Policies, Programmes and Projects

Figure 4.3 *The Green Book* Process for the Evaluation of Policies, Programmes and Projects

The second instrument is *The Magenta Book*, considered, 'The recommended central government guidance on evaluation that sets out best practice for departments to follow.'[640] Focused on evaluation, this document addresses a series of issues that public officials might face when commissioning evaluation studies (Table 4.2).[641]

---

[640] —. 2011c. "The Magenta Book. Guidance for evaluation." edited by HM Treasury. London, UK. P. 7.

[641] Henkel analyses in detail the use of consultants for the performance of evaluation studies. Henkel, M. 1991. *Government, Evaluation and Change*. London, UK: Jessica Kingsley Publishers..

Table 4.2 Key Policy Evaluation Themes Addressed by *The Magenta Book*

Theme	General Overview
Conceptualization of evaluation	Evaluation is understood as 'an objective process of understanding how a policy or other intervention was implemented, what effects it had for whom, how and why.'[642]
Evaluations approaches	The design of an evaluation depends on the kind of questions that it wants to answer; however, three types of evaluation can be identified: process evaluations, impact evaluations, and economic evaluations (cost-benefit analysis).
Planning of evaluations	It focuses on the importance of considering key elements for the performance of an evaluation such as the target audience, the approach, the resources required and appropriate timing.
Process of evaluation	The book describes a series of stages (at a more disaggregated level than those described by *The Green Book)* for the performance of an evaluation.
Performance of impact evaluation	*The Magenta Book* provides some guidance on impact evaluation, which is considered a useful tool for knowing the effect of a policy on a particular group of the population; it focuses on the performance of randomised controlled trials and piloting.
Technical issues of policy evaluation	The book addresses some technical issues regarding policy evaluation that public officers should be aware of such as data collection, reporting and dissemination of evaluation findings.

Source: Based on HM Treasury, 2011. *The Magenta Book.*

Both instruments aim to support the commissioning of evaluations under a scheme of 'objective' generation of evidence. This suggests that these processes are underpinned by an idea of rationality and depoliticisation of policymaking. The focus of these guidelines is the identification of the net effect of public policy (impact) which might explain why the idea of building counterfactuals seems so relevant. Nonetheless, this model does not appear to fit appropriately in all cases. As discussed in the case of the Cancer Strategy, the complexity of public policy exceeds this framework in the sense that it is not possible to accurately determine the impact of a policy. This means, for example, that providing access to health services (and measuring it) does not explain completely the improvement or decline in the quality of life of patients.

This rationale is also observed in the *What Works?* initiative.[643] These centres

---

[642] HM Treasury. 2011c. "The Magenta Book. Guidance for evaluation." edited by HM Treasury. London, UK.

[643] These centres have been developed to foster the production and use of evidence in social policy. Taking NICE as a point of reference, the government decided to explore 'the creation of independent expert bodies in other areas of public services.' A member of the civil service 'will be appointed as National Advisor to engage with ministers and other stakeholder groups and promote high standards across the What Works network, they will sit within the Cabinet Office and report to the Minister for Government Policy and the Chief Secretary to the Treasury.'

'will provide robust and high quality synthesis of the research evidence on the effectiveness of interventions in each field.'[644] Their technical strength is producing evidence as an aid for policymakers and politicians to decide and act under more controlled and certain conditions. At the policy level, health and education have been the areas in which the use of evaluation has been more intensive.[645] In the health sector, after the publication of the health research strategy, the DoH created the National Institute for Health Research (NIHR) 'to provide the framework through which we can position, manage and maintain the research, research staff and infrastructure of the NHS in England as a virtual national research facility.'[646] In the context of the NIHR, the Health Technology Assessment Programme (HTA) of the NHS illustrates the relevance given to evaluation in this sector. This programme funds 'independent research about the effectiveness, costs and broader impact of healthcare treatments and tests for those who plan, provide or receive care in the NHS.'[647] This salience of evaluation in health policy can be explained because 'there is a clear focus on providing evidence of efficacy or effectiveness: which technologies or other interventions are able to bring about desired outcomes for different patient groups.'[648] For this reason, it is understandable that NICE was one of the first bases of the *What Works Network,* now expanded to other areas such as crime reduction and local economic growth.[649] None of these initiatives can completely abstract the political element from policymaking. As Dorey suggests, the *What Works Network* will eventually face the complexity of the reality of the public sector in terms of the people who will prioritise the evidence to be used to hold political values that cannot be ignored.[650] In synthesis, internal evaluation in central government emphasises the linkage between public

Cabinet Office. 2013b. "What Works: evidence centres for social policy." London, UK. P. 4.

[644] Ibid. P. 2.

[645] The creation of the EEF makes evident the primary role given to evaluation in the education sector.

[646] Department of Health. 2006b. "Best Research for Best Health." London, UK.P. 9.

[647] See: National Institute for Health Research. 2015. "Health Technology Assessment (HTA) Programme." London, UK.

[648] Nutley, S., H. Davies, and I. Walter. 2002. "Evidence Based Policy and Practice: Cross Sector Lessons From the UK." in *Social Policy Research and Evaluation Conference*. Wellington, New Zealand. P. 3.

[649] NICE was established as a Non Departmental Public Body (NDPB) by the *Health and Social Care Act* 2012, its aim is 'developing guidance and quality standards in social care.' National Institute for Health and Care Excellence. 2014. "Who we are." London, UK.

[650] Dorey, P. 2014. *Policy Making in Britain: An Introduction*. London, UK: SAGE.

spending and performance and the use of indicators is an activity already embedded in the logic of departments.[651] Evaluation, in contrast, has received more attention during recent years, but it has not been internalised at the same level as performance monitoring.

A recent NAO study is consistent with the argument made here about the fragmentation of the evaluation policy. This report 'focuses on impact and cost-effectiveness evaluation relating to government spending, taxation and regulatory interventions, across the main 17 departments and some of their bodies.'[652] The analysis addressed four major issues:

1. Coverage
2. Quality
3. Use of evaluation evidence
4. Production, resources, and barriers.

The general conclusion exposes a heterogeneous panorama of evaluation in central government with gaps in terms of the quality of the studies, variety in the processes for commissioning and conducting evaluation, and very little information about the general spending devoted to the assessment of public policy and its cost-effectiveness. The Institute for Government also recognised the need for a more political approach to this activity: "One of the challenges is to move beyond the more 'technocratic' end of the policy spectrum into more 'political' or ideological areas."[653] In their report, *Making Policy Better,* this organisation highlights:

> While government often commissions evaluations, our evidence shows that most politicians and civil servants are extremely sceptical about whether Whitehall takes note of their results: lessons often do not feed back into policy design or problem formulation [...] although evaluations are often commissioned they are often ignored.
>
> Evaluations are usually commissioned and managed by the same department that carried out the policy [...] the department has the incentive and opportunity to tone down evaluation findings that are critical, but which could lead to significant learning, since evaluators often depend on repeat business, they have the incentive to acquiesce in self-

---

[651] See for example: HM Treasury. 2013. "Spending Round 2013." London, UK.

[652] National Audit Office. 2013a. "Evaluation in government." edited by National Audit Office. London, UK. P. 6.

[653] Rutter, J. 2012. "Evidence and Evaluation in Policy Making. A problem of supply or demand?", edited by Institute for Government. P. 29.

censorship.[654]

This last finding can be easily observed in practice. In the health sector, for example, there are some consultancy firms who are recurrent providers of evaluation studies for different public bodies. Ipsos Mori,[655] for instance, during the last year produced diverse studies for the DoH, the NHS, and PHE,[656] and this also occurs in other sectors, e.g. Coffey International Development for the DfID.[657] As departments require evidence to support the existence, continuity, or elimination of their programmes, there are incentives for commissioners and evaluators to establish informal arrangements for the pursuit of individual interests, e.g. obtaining contracts, or producing more favourable evidence.

Moreover, the lack of information about the overall situation of policy evaluation in central government hinders the understanding of the extent to which this activity has penetrated into the public sector logic and the degree to which it has been able to enter into the political system. The studies undertaken by both the NAO and the Institute for Government provide some insight about the politicisation of this activity that will be reviewed later in Chapter 6. The next section addresses the external audit system in the UK.

4.1.3 External Auditing

Auditing in the UK is an activity embedded in the public sector at different levels. Although other audit institutions exist in the UK, this section focuses on the national level; specifically on the work of the NAO (external evaluation).[658]

---

[654] Rutter, J., and M. Hallsworth. 2011. "Making Policy Better. Improving Whitehall's core business." London, UK: Institute for Government. P. 32.

[655] Ipsos Mori is one of the most well-known consultancy firms in the UK in the field of social and political research. See: Ipsos Mori. 2015. "About us." London, UK.

[656] Some examples are the *Cognitive testing of integrated care questions* commissioned by the Health and Social Care Information Centre (HSCIC) and the Department of Health; the *Public Health England 2014 public awareness and opinion survey* commissioned by PHE; and the *Friends and Family Test Qualitative Research* on behalf of NHS England.

[657] For example, the study commissioned about small and medium-sized enterprises in Afghanistan. See: Coffey International Development. 2014. "Scoping the Need for Improved Access to Funding, Advisory Support and Business Skills Development for SMEs in Afghanistan." London, UK.

[658] Until the 31st March 2015 the Audit Commission was in charge of scrutinizing local services in England, currently a new set of procedures is taking place in which the functions of this body are now distributed among the Public Sector Audit Appointments Limited, the NAO, the Financial Reporting Council and the Cabinet Office. See: Department for Communities and Local Government. 2015b. "Local audit framework replacing the Audit Commission." London, UK.

This body is in charge of inspecting public spending on behalf of the Parliament[659] through two major activities:

> Audit financial statements of all central government departments, agencies and other public bodies and report the results to Parliament, and
> Value for money studies[660] to look at how government projects, programmes and initiatives have been implemented; they contain recommendations on how services can be improved.[661]

Box 4.2 An Overview of the NAO

> The NAO in its present form was established in 1983 as a corporate entity with a statutory Board; however, the UK has a long tradition in the auditing field and this body has evolved along time. There are records dated back to 1314 that mention a public official charged with auditing government expenditure. The *Budget Responsibility and National Audit Act 2011* set the basis for the governance of the NAO:
>> It requires that the Comptroller and Auditor General (C&GA) and Chairman are to be appointed by Her Majesty the Queen, upon an address to Parliament by the Prime Minister.
>> The Public Accounts Commission is formally responsible for the appointment of the non-executive members of the NAO Board and the external auditor of the NAO.
> Its budget is approved by the Public Accounts Commission, who also evaluates its performance and it can suggest the commissioning of VFM studies. For the 2013-2014 period, the NAO has 822 permanent members of staff and a £63 million budget for its operation.

The NAO is regulated by two major Acts of Parliament: the *National Audit Act 1983* and the *Government Resources and Accounts Act 2000* (Table 4.3).[662]

---

[659] As a parliamentary body, the NAO is outside the government's sphere. This is a common feature of audit bodies linked to the Legislative, for example, the Government Accounting Office that reports directly to the U.S. Congress or the Supreme Audit Office in Mexico linked to the Chamber of Deputies.

[660] A value for money study 'focuses on a specific area of government expenditure, and seeks to reach a judgement on whether value for money has been achieved, we define good value for money as the optimal use of resources to achieve the intended outcomes.' See: National Audit Office. "What is a Value for Money Study?". London, UK.

[661] —. 2013k. "An Introduction to the National Audit Office." London, UK.

[662] *Appendix E* presents a synthesis of the history and development of the NAO prior to its consolidation in 1983.

Table 4.3 Legislation Underpinning the Work of the NAO

Act of Parliament	Key issues
National Audit Act 1983	• It establishes the National Audit Office and the Public Accounts Commission as the bodies in charge of controlling public expenditure on behalf of Parliament • It names the C&GA the head of the NAO and it gives him 'complete discretion in the discharge of his functions' (Part I, section 3). • The C&GA 'may carry out examinations into the economy, efficiency and effectiveness with which any department, authority or other body [...] has used its resources in discharging its functions' (Part II, section 6). • The NAO reports to the Public Accounts Commission.[663]

Government Resources and Accounts Act 2000	• It is an Act 'make provision about government's resources and accounts...'
	• It establishes that government departments shall prepare accounts for every financial year.
	• Departments must heed the directions given by the Treasury for the preparation of accounts.
	• These accounts must be sent to the Comptroller and General Auditor.
	• The C&GA must examine the accounts received and verify:
	    the accounts present a true and fair view,
	    that money provided by Parliament has been expended for the purposes intended by Parliament,
	    that resources authorised by Parliament to be used have been used for the purposes in relation to which the use was authorised, and that the department's financial transactions are in accordance with any relevant authority (c. 6).
	• The C&GA must issue a report of the examination of the accounts.

There are five key areas in the NAO's work: financial audit, VFM studies, insight and investigations, support to Parliament, and international audit.[664] Its tasks are more related to accountability than to evaluation because the core of its activities is verifying the appropriate use of public funds, and through feed into an accountability process (Figure 4.4).

---

[663] The *National Audit Act 1983* and the *Budget Responsibility and National Audit Act 2011* establish the duties of the Public Accounts Commission, some of the most relevant are: 'to examine the National Audit Office Estimate and lay it before the House, to consider reports from the appointed auditor of the National Audit Office, to appoint non-executive members of the National Audit Office Board, and to report from time to time.' See: The Public Accounts Commission. 2012. "Role." London, UK.

[664] National Audit Office. 2013k. "An Introduction to the National Audit Office." London, UK.

Figure 4.4 The Accountability Process
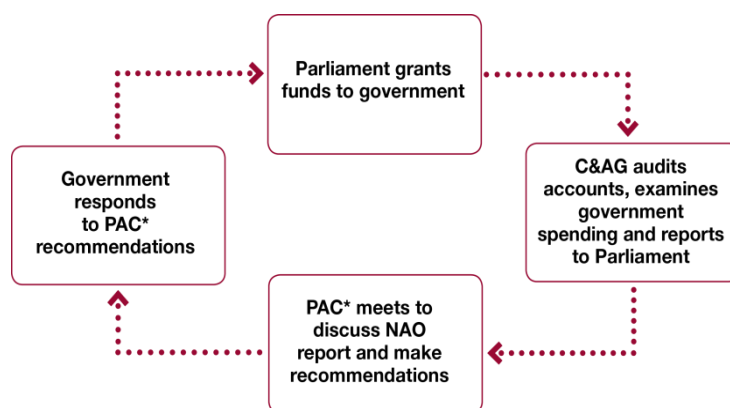
Source: National Audit Office, 2011. *An Introduction to the National Audit Office*. P. 7.

In terms of evaluation, the work of the NAO is relevant for the performance of VFM studies that generate recommendations for the improvement of public services (Figure 4.5). During 2013/14, for instance, the NAO delivered 66 VFM reports covering an important range of policy areas.[665] These studies identified if government departments were delivering services effectively and provided a series of recommendations to improve VFM.[666] As the NAO does not have faculties for questioning the value or merit of policy, these studies aim 'to form a clear judgement on whether value for money has been secured.'[667] Through the analysis of the outputs of policy, VFM's studies …

> … consist of detailed examinations of specific spending programmes within the relevant department in order to weigh up how efficiently, effectively and economically the government has used public money to achieve its aims.[668]

This has produced two major consequences:

> A much greater focus on outcome achievement and the application of a much wider and sophisticated range of diagnostic and analytical techniques underpinning assessments of value for money, and consequently a need to rely on a much more multidisciplinary skill set.[669]

---

[665] —. 2014a. "National Audit Office. Annual Report and Accounts 2013 -14." edited by National Audit Office. London, UK.

[666] See for example: —. 2012. "The management of adult diabetes services in the NHS." London, UK.

[667] —. 2013n. "National Audit Office. Annual Report and Accounts 2012-13." edited by National Audit Office. London, UK. P. 1.

[668] Bacon, R., and C. Hope. 2013. *Conundrum. Why every government gets things wrong -and what we can do about it*. Great Britain: Biteback Publishing Ltd. P. xvi.

[669] Bourn, J. 2007. *Public Sector Auditing: Is it value for money?* Chichester, UK: John Wiley &

Auditing is related to values like objectivity and independence. However, as Gray and Jenkins recognise, 'In the world of VFM audit […] studies are conducted in an arena of greater uncertainty and political sensitivity.'[670] This reflects the complexity of separating activities like auditing and evaluation. The major difference that arises is the nature of the analytical instruments and the scope selected.

Figure 4.5 The VFM Study Cycle
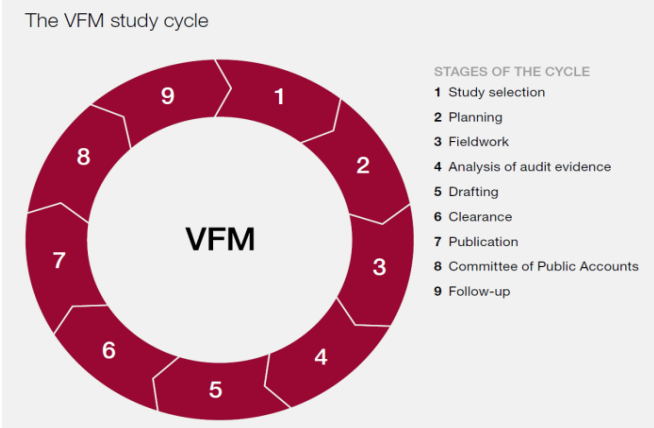
Source: National Audit Office. Annual Report and Accounts 2012-13. P.15.

VFM studies do not question the government's decisions on policy, but they do bring attention to multiple policy issues in the political agenda by providing evidence to Parliament about how money has been spent and the improvement of public services. The VFM programme is developed through an internal process in which NAO senior management decides its content, but that also needs to be reactive to the agenda of the PAC, as an NAO official explains (UK-3):

> We have a good understanding about the on-going issues, the main issues that we want to focus across the year, so we would put forward a range of proposals […] and then senior management would decide which proposals they would like to take forward […] We may have to prioritise work so if an issue comes up that we need to look at quickly, a piece of work that we were due to do may get pushed back, so that we can fit in and be more reactive, so it's quite a sort of fluid process …

This reveals that although there are rules and processes in place for developing

Sons, Ltd. P. 58.

[670] Gray, A. , and B. Jenkins. 1993. "Separate Developments: Budgeting, Auditing, and Evaluation in the United Kingdom." in *Budgeting, Auditing, and Evaluation. Functions and Integration in Seven Governments* edited by A. Gray, B. Jenkins, and B. Segsworth. London, UK: Transaction Publishers.

the VFM programme, this is subject to the discretion of the actors involved in the definition of the priority issues in the agenda.[671]

The differences between evaluation and auditing form a point that requires more attention. Considering the work of the NAO as evaluation is debatable for two reasons. First, because appraising the effectiveness of the government in delivering public services does not necessarily imply the identification of a policy's value. Secondly, because although they do not judge the rationale behind a programme, they do assess a set of features of policy that lead to recommendations, which can eventually be transformed into decisions. Consequently, the boundaries between judging the implementation of a policy and judging the policy itself are somewhat blurred. Information about the efficiency of a particular programme or the economy of a strategy can represent an important asset for decision-making, such as budget allocation or the prioritisation of policy issues. Also, it can represent an opportunity for the opposition to evidence the failures of government.[672] An external view on the effectiveness of their policies can generate evidence for government legitimisation.[673] These differences can therefore be observed in the rationale of these activities (control vs improvement/learning), the methods applied (financial tools vs social research methods), and the focus of analysis (deviations vs results/impacts).

In synthesis, in the UK it is possible to identify two main strands for the evaluation and audit of public policy. On the one hand is performance monitoring (focusing on the link between public spending and results) in addition to evaluation studies about different dimensions of policy. HM

---

[671] This occurred in the CQC inquiry, 'the CQC was an example of where the PAC asked us to do some work because at the time there had been a BBC Panorama Programme on something called Winterbourne View and also one of the big providers, Southern Cross, was having some difficulties' (UK-3).

[672] Greener, I., and B. Greve (Eds.). 2014. *Evidence and Evaluation in Social Policy*. London, UK: Wiley Blackwell.

[673] One of the major attributes associated to the NAO is its independence, 'the constitutional framework of the NAO is intended to provide the degree of independence necessary to allow public sector audit to effectively contribute to the accountability of government.' This becomes of great relevance when comparing the role of the NAO with its Mexican homologue in the analytical chapter of this thesis. White, F. 1999. *Audit, Accountability and Government*. Oxford, UK: Clarendon Press. P. 95.

Treasury can be identified as the leading organisation that controls the performance monitoring system. However, as there is no formal evaluation policy within central government, the development of this activity depends on the rules of each organisation.

On the other hand, the NAO concentrates on scrutinising the government's capability to deliver effectively public services, but this body does not formally question the pertinence of the decisions made to address those problems.[674] Beyond the utility of valuing the effectiveness of public services, the most important contribution of the NAO relies on the legitimacy and trust of its work.[675]

As the major purpose of this research is to identify the relationship between politics and evaluation, it seems important to make the following remarks about the UK case:

1. Performance monitoring is an institutionalised practice that produces a bond between public spending and performance. Central government departments have internalised the monitoring of their activities through the measurement of a set of indicators (e.g. the number of overseas visitors to the UK.)[676]
2. The lack of an integrated evaluation policy makes difficult the identification of formal guidelines about the independence of evaluations. The DfID, for example, has developed its own mechanisms for preserving the independence of their studies, but there is no common procedure for ensuring the independence of all evaluation studies.[677]
3. In relation to the studies commissioned to external evaluators, the criteria for commissioning the study, as well as the interaction between the evaluator and the department are interesting elements of study for identifying the political struggles that might take place. The process for the allocation of public contracts is underpinned by the logics of transparency. However, the explicit criteria established for assessing a bidder's proposal can be subject to the discretion of the people in charge of the decision.[678]

---

[674] For further reference about the role of PAC see: Dunleavy, P., C. Gilson, S. Bastow, and J. Tinkler. 2009. "The National Audit Office, the Public Accounts Committee and the risk landscape in UK public policy." London, UK: The Risk and Regulation Advisory Council.

[675] Talbot, C., and J. Wiggan. 2010. "The public value of the National Audit Office." *International Journal of Public Sector Management* 23(1):54-70.

[676] Impact indicator. Department for Culture, Media and Sport. July 2013.

[677] The work of the DFIF is supported by the Independent Commission for Aid Impact (ICAI), which 'is the independent body responsible for scrutiny of UK aid. ICAI focuses on maximising the impact and effectiveness of the UK aid budget for intended beneficiaries and the delivery of value for money for the UK taxpayer.' See: Independent Commission for Aid Impact. 2015. "ICAI." London, UK.

[678] The set of criteria for evaluating a tenderer's proposal are clearly stated; however, it demands the interpretation of the commissioner which implies a degree of discretion. See for example: Contracts Finder. 2015. "The Provission of A & E Qualitative Research." London, UK.

4. Due to the fact that evaluation capabilities are uneven within central government, it is relevant to analyse whether there are areas in which the political commitment to evaluation is stronger; the reasons behind this situation. The DfID, for example, is outstanding from the rest of the government departments because of its explicit commitment to evaluation.[679] The asymmetry in the level of maturity in terms of evaluation can be explained by multiple factors. e.g. the participation of non-governmental stakeholders.[680]

5. The relationship between the NAO and central government departments can interfere with independence and objectivity,[681] but at the same time it can represent a mechanism to reduce political confrontation.[682]

This section has shown that in the UK case, internal and external processes of evaluation operate simultaneously, but in a disconnected way. This matters for the study of the hidden politics of evaluation because it suggests that there are spaces of discretion that stakeholders can take advantage of and that can produce asymmetry of power and resources between them. The following sections describe the two case studies analysed in the UK. The first one is the Cancer Strategy implemented by the DoH, while the second case analyses the CQC, a body that regulates the provision of health and adult social-care services. The aim has been to examine the two evaluation streams and to expose the hidden politics of evaluation.

---

[679] This department has an explicit evaluation policy, it systematically publishes evaluation reports and it has been recognized by international organisations like the OECD and The World Bank for its efforts in the area of monitoring and evaluation. See: York, N. 2013. "Developing an Evaluation Strategy-experience in DFID."

[680] An example of this situation is the evaluation of the programme for Strengthening South Africa's Revitalised Response to HIV and Health, it involved the participation of the National Department of Health (South Africa), Mott MacDonald Ltd and the DFID, among others. See: Griffiths, S., S. Ahmar, J. Seager, L. Simbayi, and N. Mbelle. 2014. "Mid-Term Evaluation of the SARRAH programme." Department for International Development.

[681] Dunleavy et al. state that 'departments' amendments often serve to 'blunt' NAO's criticism, creating blander and sometimes euphemistic language." Dunleavy, P., C. Gilson, S. Bastow, and J. Tinkler. 2009. "The National Audit Office, the Public Accounts Committee and the risk landscape in UK public policy." London, UK: The Risk and Regulation Advisory Council. P. 4.

[682] For example, as it is shown in Figure 4.5, one if the stages of the VFM cycle is *clearance,* which can represent a threaten to the independence of the study or compromise the NAO to modify their original remarks on a policy, this will be discussed with more detail in subsequent sections of this chapter.

4.2 Case A: The Cancer Strategy

Understanding the relationship between evaluation and politics cannot be limited to the rules established by the formal institutional framework. Instead it is necessary to reach a deeper level of analysis through the description of how evaluation is performed in practical terms. This section is dedicated to the case of the Cancer Strategy, a policy implemented by central government to tackle a disease that will affect '1 in 2 people in the UK in their lifetime'.[683] As a response, in 2011 the Coalition Government published *Improving Outcomes: A Strategy for Cancer* (IOSC)*,* a policy document that guides governmental action in this matter:

- It sets out the actions we will take to tackle the preventable causes of cancer, by providing better information to people about risk factors and how individuals and communities might work to minimise them, as well as the steps we will continue to take to improve the experience of cancer patients and support the increasing number of cancer survivors.
- Describes the ways in which choice for patients in their cancer care will be extended and implemented throughout the health and social-care systems, informing both the decisions taken by NHS organisations now and the methods through which the mandate for the NHS Commissioning Board may be discharged; and
- Identifies the gaps in information on health outcomes which are crucial to ensuring patients are empowered – in consultation and with the support of their clinicians – to exercise real choice over the care they receive, including through the extension of national clinical audit and through the strengthened patient voice delivered by Health Watch.[684]

The Cancer Strategy represents a useful example of a policy area in which there has been a systematic production of knowledge, both generated by governmental bodies such as the National Cancer Intelligence Network (NCIN) and NICE, as well as by an active community of NGOs actively engaged in the provision of evidence about a disease for which 'incidence rates in Great Britain have increased by more than a third since the mid-1970s.'[685]

This case study describes the evaluation and auditing processes of the Cancer Strategy to set the basis for the analysis of the political relationships that are built as a consequence. The core argument presented here is that a policy issue

---

[683] Cancer Research UK. 2015e. "What is cancer?". London, UK.
[684] Department of Health. 2011b. "Improving Outcomes: A Strategy for Cancer." London, UK. P. 3.
[685] Cancer Research UK. 2013. "Cancer Statistics Key Facts." London, UK.

that heavily relies on scientific evidence and knowledge is less susceptible to the politicisation of evaluation. However, in the case of the cancer policy the sensitiveness of this issue inevitably entails an important amount of public attention that guarantees its permanent position in the political agenda. Cancer is a recurrent subject on the parliamentary agenda and there is a careful follow-up of the progress in this matter. For example, as the PAC's Chair recently stated:

> It is good news that survival rates for cancer patients are improving with 69% of cancer patients now surviving 1 year after diagnosis and 49% surviving 5 years after diagnosis. Yet, I am deeply disappointed that cancer survival rates in England remain poor compared with the best European countries, with rates in England still 10% lower than the European average according to the latest data. This is a shocking disparity.[686]

Importantly the relevance of cancer as a disease that affects the British population is reflected both in the design and implementation of different policy initiatives, as well as in the level of scrutiny applied to this policy. This section provides an overview of the Cancer Strategy, about the tools that have been applied for its monitoring and evaluation and the audit process implemented.

4.2.1 General Overview of the Cancer Strategy

According to Cancer Research UK, 'Cancer is a major cause of mortality and morbidity in the UK [...] more than one in four (29%) deaths were caused by cancer in the UK in 2011.'[687] This has inevitably forced the government to take action in the development of policies that can improve the conditions of the population suffering this disease. Different actions have been directed to make and measure the progress made by policy in health outcomes. This has been adopted as a political priority: 'We will measure our success on the health results that really matter – such as improving cancer and stroke survival rates or reducing hospital infections.'[688] Specifically, there has been an increasing concern about a better allocation of resources and in the improvement in the

---

[686] Committee of Public Accounts. 2015a. "Chair's statement on improving cancer services and outcomes in England."
[687] Cancer Research UK. 2014a. "Cancer Statistics Report: Cancer Mortality in the UK in 2012." P. 1.
[688] HM Government. 2010. "The Coalition: Our programme for government." London, UK. P. 25.

quality of the services delivered to patients.[689]

The Cancer Strategy was the most representative instrument of the Coalition Government (2010–15) to tackle this disease. It is a policy instrument, the implementation of which involves the DoH and other organisations such as the NHS and PHE (Figure 4.6). The design, operation, and evaluation of the Cancer Strategy implies a complex system of stakeholders that need to reach agreements about the feasibility of the strategy and for the achievement of organisational goals. As a DoH official (UK-8) explains:

> Our strategy is supported by an impact assessment that had to be signed off by analysts and then [...] by a Government Minister, saying that all things in the strategy [that] would cost money were evidence-based, were cost-effective and were affordable ...

This finds echo in the discussion about coordination in governance (see Chapter 2, Section 2.1.1). The interaction of multiple actors in policymaking can lead to the rise of conflicts that can affect implementation and evaluation or, as Bovens defined, *a problem of many hands.* The achievement of targets in the tackling of cancer depends not only on the development of key strategies by the DoH, but also on the effective provision of services to patients by the NHS and the development of evidence about the disease by PHE, among other elements. This case makes evident the need for establishing mechanisms to cooperate in the pursuit of a goal, e.g. the reduction of mortality rates.

---

[689] See for example: Department of Health. 2015b. "Safer care for patients." London, UK.

Figure 4.6 Department of Health (Agencies and Public Bodies)

Executive agency		
Medicines and Healthcare Products Regulatory Agency	Public Health England	
Executive non-departmental public body		
Care Quality Commission	Health and Social Care Information Centre	Human Fertilisation and Embryology Authority
Human Tissue Authority	Monitor	NHS England
National Institute for Health and Care Excellence		
Advisory non-departmental public body		
Administration of Radioactive Substances Advisory Committee	Advisory Committee on Clinical Excellence Awards	British Pharmacopoeia Commission
Commission on Human Medicines	Committee on Mutagenicity of Chemicals in Food, Consumer Products and the Environment	Independent Reconfiguration Panel
NHS Pay Review Body	Review Body on Doctors' and Dentists' Remuneration	
Other		
Accelerated Access Review	Health Education England	Health Research Authority
Morecambe Bay Investigation	NHS Blood and Transplant	NHS Business Services Authority
NHS Litigation Authority	NHS Trust Development Authority	National Information Board

Source: gov.uk

The DoH's performance, in both cases analysed in the UK, relies on the capacity of other bodies to comply effectively with its tasks. In the case of the Cancer Strategy, for instance, the collection of data about cancer services (by the NHS) is a requirement for the development of research in this field, e.g. by NICE. Consequently, the existence of multiple decision points, the asymmetry of resources between these bodies, and the political will to cooperate are only some of the elements that need to be considered when looking at the performance of the DoH. This complexity cannot be ignored in evaluation—even less in its accountability functions—because it increases the difficulty of assigning responsibilities for the success of a policy. This provides a clear example of the transition from government to governance. The DoH must operate within a structure in which different networks are established and when hierarchical relationships become obsolete.

Since the year 2000, the governmental response to cancer has transited through different reforms that eventually led to what is known as the Cancer Strategy.[690]

---

[690] Under different names and models, the central government has implemented different measures for tackling the cancer problem in the UK. The *NHS Cancer Plan: A plan for investment, a plan for reform* published in 2000 can be considered one of the foundations of the *Cancer*

177

This has been mostly produced by the different political shifts in central government and by the pressures generated by the evolution of this disease among the British population. The arrival of the Coalition Government meant the revision of the strategy that …

> … set out how the Coalition Government's commitments to enable patients to rate hospitals and doctors according to their quality of care, and putting patients in charge of making decisions about their care, will be achieved for cancer.[691]

In 2011, this revision led to the publication of the document, *Improving Outcomes: A Strategy for Cancer*, which established a series of actions and strategies that made explicit the government policy in this area (Box 4.3).

Box 4.3 The Cancer Strategy

The current *Cancer Strategy* derives from the *Cancer Reform Strategy* (CRS) from 2007 which established a series of actions directed to improve cancer outcomes and to ensure the delivery of services. The CRS incorporated actions directed to cancer prevention, earlier diagnosis, reducing cancer inequality, among others. The ***Improving Outcomes: A Strategy for Cancer*** document presents the reformulation of the strategy by establishing three principles to guide the reforms undertaken in the health sector (p.2):

1. To put the patient or service user at the heart of the public services –transforming the relationship between citizen and service through the principle of *no decision about me without me*;
2. To orientate the NHS, public health and social care services towards delivering the improvements in outcomes which matter – rather than measuring processes which do not; and
3. To empower local organisations and professionals to deliver the freedoms to innovate and to drive improvements in services which deliver care of the highest quality for all patients and service users.

This strategy pursues three objectives: 1) to generate more useful and relevant information for 'patients, commissioners, providers, clinicians, researchers' (p.18) to make more informed decisions, 2) to improve cancer outcomes that can be monitored by specific indicators; and 3) to reduce cancer inequalities among the different groups of population in the country.

Source: Based on Department of Health, 2011. *Improving Outcomes: A Strategy for Cancer.*

The Cancer Strategy covers a vast range of activities guided by the principles established in the strategy. This increases the complexity of grasping not only

---

Reform Strategy published in 2007 and that guided the publication of *Improving Outcomes: A Strategy for Cancer* in 2011.

[691] Richards, M. 2010. "Review of the Cancer Reform Strategy." edited by Department of Health. London, UK.

how the strategy is implemented, but also how these activities can be evaluated and their contribution to its 'success' or 'failure'. Moreover, the different organisations involved in the execution of the strategy and the conceptualisation of its functioning increases the difficulty of pointing out the responsibility of the actors involved.

The systematic use of evidence in this case is associated with the recognition of those areas in which there can be improvement in efficiency and effectiveness. The director of an NGO (UK-11) highlights the relevance of the strategy by stating that this 'is looking at key aspects that the health service can have a very big impact on'. This concerns the discussion about seeking for more efficiency as one of the drivers of the transformation of the state. Those policy areas in which more effects can be perceived can become a priority in terms of the proportion of the population that is being served. In this case, the focus on outcomes has aimed to improve effectiveness. To illustrate, earlier diagnosis was evaluated in terms of its impact on cost and benefits for the NHS:

> Earlier diagnosis is generally cost effective, but not cost-saving. If people are diagnosed earlier, either through screening programmes or through their general practice, the main benefit is a substantial improvement in health outcomes. There is not a cost reduction, rather an increase in NHS costs (large increase in testing costs generally offset by a modest reduction in treatment costs).[692]

The Cancer Strategy implied also the reallocation of public funds to achieve the goals established in the IOSC. The implementation of this strategy represented additional funding of more than £750 million in a four-year period.[693] This reflects the magnitude of cancer in the UK, not only in the health policy but also in terms of its economic impact. As the NHS recognises, 'Annual NHS costs for cancer services are £5 billion, but the cost to society as a whole – including costs for loss of productivity – is £18.3 billion.'[694] Even though most of the actions are consistent with the efforts made by governments in the past e.g. earlier diagnosis, changes in lifestyle, etc., the Coalition Government emphasised the need to focus on outcomes that can be assessed:

---

[692] Department of Health. 2011f. "The likely impact of earlier diagnosis of cancer on costs and benefits to the NHS." London,UK. P. 22.
[693] National Health Service. 2014. "Improving outcomes: a strategy for cancer." London, UK.
[694] Department of Health. 2013c. "Helping more people survive cancer." London, UK.

> Governments of the past have placed too much faith in the ability of a top-down hierarchy to deliver improved results for patients, through the micro-management of services from Whitehall resulting in the disempowerment of frontline staff and the disenfranchisement of patients. Too little faith has been placed in the power of local communities and frontline clinicians – and patients and service users themselves – to drive the improvements that we need to see.[695]

As a result, the evaluation of the strategy, described in the next section, is strongly linked to the values of efficiency, efficacy, and effectiveness of the outcomes set out in the IOSC.

4.2.2 The Evaluation of the Cancer Strategy

Despite the fact that there is no formal central evaluation policy, the Cancer Strategy has been systematically evaluated since 2011. It has been subject to different internal and external mechanisms of scrutiny and the production and use of research to inform policy decisions is a common practice. The DoH evaluates annually the progress of the outcomes established in the IOSC, focusing on 'the cost-effectiveness of interventions' (UK-8). In addition, it has carried out multiple evaluation studies, (both internal and commissioned to external parties) in different areas related to cancer.[696] In December 2012, for example, the DoH published a study that described the best practice …

> … in relation to piloting of service delivery to older people with cancer to improve outcomes; to deliver high quality services to increasing number of older cancer patients and to ensure quality of access to treatment and information, based on need, not age.[697]

This suggests that evidence is considered an important input for internal decision-making. Prior to the publication of the IOSC document, the DoH carried out an impact assessment to identify the best policy option for improving the outcomes of the Cancer Strategy. Before making the final decision, other policy options were considered, in line with the guidance of *The Green Book.* This

---

[695] —. 2011b. "Improving Outcomes: A Strategy for Cancer." London, UK.. P. 3.
[696] An example of these studies concerns the quality of life of cancer patients. See: Department of Health, and NHS England. 2012. "Quality of Life of Cancer Survivors in England. Report on a pilot survey using Patient Reported Outcome Measures (PROMS)." edited by Department of Health. London, UK.
[697] McMillan Cancer Support, Department of Health, and Cancer Research UK. 2012. "Cancer Services Coming of Age: Learning from the Improving Cancer Treatment Assessment and Support for Older People Project." London, UK. P. 2.

instrument offers policymakers a framework for reflecting on the need for government intervention, the objectives, and the intended effects of a particular policy, in addition to cost-benefit analysis of a potential policy.[698]

This impact assessment set out three policy options that could have been implemented to achieve better results, along with a cost-benefit analysis for each option. This evidence supported the reform of the Cancer Strategy by showing that the option chosen was the one that will produce more benefits in the long term. This point is important because it suggests that policy decisions are made in terms of the benefits that will be produced. The core element of this reflects on who defines what benefit is and how is this operationalised. So far, the progress of the strategy has been reported in four reports. Due to structural changes in the health sector, the two most recent were published by the DoH jointly with PHE[699] and NHS England.[700] Each one has focused on the areas established as priorities in the IOSC, such as early diagnosis and data-collection. Data emphases the implementation of the IOSC and the achievements made during the year. Overall, the four reports follow the same structure concentrating on key aspects of the IOSC that are evaluated in relation to previous years. This is the most evident formal mechanism of internal evaluation that can be identified in this case (Table 4.4).

---

[698] Department for Business, Innovation and Skills. 2014. "Impact assessment template for government policies." London, UK.

[699] The creation of PHE was driven by the reforms to the NHS derived from the *Health and Social Care Act 2012*. This set of reforms gave PHE different responsibilities regarding the promotion of public health, among which outstands 'reporting on improvements in the public's health so everyone can understand the challenge and the next steps.' Its inclusion as one of the actors in charge of updating the public about the progress in the Cancer Strategy represents a new model of co-responsibility that will be analysed with more detail in Chapter 6. Public Health England. 2014a. "About us." London, UK.

[700] The involvement of the NHS is an important issue to highlight. Especially in the Cancer Strategy case, the assessment heavily relies on data produced and managed by the NHS, for example the collection of data about treatment, e.g. radiotherapy, depends on the effective provision of data from the NHS. As it will be discussed in Chapter 6, data-management can represent a major resource for political control in the performance of evaluations or audits.

Table 4.4 Progress Reports on the Cancer Strategy

Year	Report	Main Findings
2011	Improving Outcomes: A Strategy for Cancer First Annual Report	This report recognised different challenges to address in relation to cancer, but some important achievements were made after the first year of implementation: there have been further improvements in data and analyses that we have to support clinicians, commissioners and providers to improve services and patients to make informed choices; we are continuing to extend the age range for bowel and breast cancer screening; we have successfully run a range of campaigns to improve the public's awareness of cancer symptoms; the NHS has continued to improve care pathways in order to reduce inpatient bed days- reducing costs and improving the quality of care for patients (pp. 3-4).
2012	Improving Outcomes: A Strategy for Cancer Second Annual Report	At this point there was still no data available to measure the progress in the ambition of saving 5,000 lives per year by 2014-2015; it also provided information about some other areas of action in which achievements can be more easily identified: there have been improvements in survival rates; the age range for the breast and bowel screening program, campaigns are running to encourage people to go to their doctor if relevant symptoms persist (in order to get the cancer diagnosed when it is more treatable) and patient's access to radiotherapy continues to rise (p.7).
2013	Improving Outcomes: A Strategy for Cancer Third Annual Report	This report is consistent with the achievements identified in previous years; yet there was still no information about the ambition of saving 5,000 lives per year by 2014-2015, it reported that 'cancer survival estimates and mortality rates continue to improve', the most relevant achievements were: significant developments in cancer screening –particularly on the first phase of introducing Bowel Scope Screening (BSS); activity to promote earlier diagnosis of symptomatic cancers, through the Be Clear on Cancer campaigns and the associated work with primary and secondary care; progress in ensuring better access for all to the best possible treatment, for example through improved access to Intensity Modulated Radiotherapy (IMRT); significant developments in the collection and reporting of new datasets and the analysis of information, to drive improvements and to inform patients (p. 6).
2014	Improving Outcomes: A Strategy for Cancer Fourth Annual Report	For the first time the DoH was able to provide new evidence about the government's ambition to save an additional 5,000 lives a year by 2014/15: Using the latest survival estimates and assuming incidence is constant, we estimate that on average between 6,500 and 17,000 more patients per year diagnosed from 2011-2015 will survive cancer for five years compared to those diagnosed from 2006-2010 [...] we also would not claim that all the improvement is down to actions set out in the IOSC, but we believe that actions set out in that Strategy have had a significant impact on cancer survival in this country (p. 8).

Source: Based on Department of Health, 2011-2014. *Improving Outcomes: A Strategy for Cancer.* Annual Reports.

The progress observed in different areas of the IOSC is plausible due to the time in which the strategy has been implemented. This includes improvements in the Cancer Patient Survey, developments in the collection and reporting of data, access to radiotherapy treatment, among others.[701] The late diagnosis and its effect on the efficiency of the governmental action are still debatable issues. According to *The Guardian* and Cancer Research UK:

> Almost half of people who get cancer are diagnosed late, which makes treatment less likely to succeed and reduces their chances of survival [...] Urgent improvements to how soon cancer is diagnosed would save the NHS tens of millions of pounds a year through reduced chemotherapy, radiotherapy and surgery, as well as enhancing many cancer sufferers' chances of survival, CRUK estimates.[702]

The systematic evaluation of the Cancer Strategy has been mostly constrained to the identification of progress in a group of variables and its behaviour in relation to efficiency, efficacy, and effectiveness, and yet, multiple variables associated with cancer cannot be addressed in deepness in the DoH reports. For example, the relationship between smoking and cancer is an issue that has been debated both within and outside the governmental sphere, but the level of specificity about this topic exceeds the scope of the DoH reports. [703] This will be discussed in more detail in the analytical chapter.

In terms of external scrutiny, in 2010 the NAO produced a report[704] on the Cancer Strategy that aimed to examine ...

> ... how effectively three of the Strategy's actions to drive delivery have been utilised to improve services for cancer patients: 1) improving the quality of information, 2) strengthening commissioning, and 3) making better use of resources.[705]

---

[701] Department of Health, Public Health England, and NHS England. 2014. "Improving Outcomes: A Strategy for Cancer. Fourth Annual Report." London, UK.

[702] Campbell, D. 2014a. "Almost half of cancer patients diagnosed too late." in *The Guardian*. London, UK.

[703] For example, the discussion about cigarettes packaging has become a central issue because of its relation to cancer in the UK. See: Department of Health. 2015a. "Government backs standarised packaging of tobacco." Cancer Research UK. 2015d. "Standard cigarette packs 'could save 2000 lives a year'." London, UK.

[704] It is important to state that the report delivered by the NAO was based on the Cancer Strategy Reform published in 2007. The current operating strategy is the one reformed in 2011. *Appendix G* presents a general description of the 2007 CRS.

[705] National Audit Office. 2010. "Delivering the Cancer Reform Strategy." London, UK. P. 5.

The findings derived from this study focused on the DoH's capacity to allocate its resources efficiently.[706] Mostly, it identified those areas in which it is possible to reduce costs and to achieve better outcomes. The report emphasised the opacity in costs of the services delivered by commissioners, which has led to imprecisions in the appraisal of the cost-effectiveness of the strategy, and it manifests a fragmentation in the cancer service delivery.[707] Nonetheless, as an NAO member recognises (UK-4) this was 'quite a positive report recognising the progress':

> The Department has made progress in improving key aspects of cancer services through strong direction and high profile leadership underpinned by increased resources. While there have been measurable improvements in efficiency by treating more people as day cases, and reductions in length of stay, there is substantial scope to make further improvements by tackling variations and raising performance to the standard of the best.[708]

One element to highlight was the role of the then Cancer Director, Professor Mike Richards, and how he championed the use of evidence: 'He ran cancer in the UK and he ran it regardless who was in government […] he believed in data and […] in being analytical' (UK-11). Not only the strength of evidence has been a key factor in the implementation and evaluation of the Cancer Strategy. The leverage given to this policy by key stakeholders has also promoted underlined the relevance of guiding decision-making on robust data.

More recently, the NAO reviewed the progress in improving cancer services and outcomes in England. The report (January 2015) examines the implementation of the recommendations made by the PAC in 2011. Some of the key findings of this study are summarised below:

- Outcomes for cancer patients, including estimated survival rates, continue to improve.
- However, there remains considerable scope to improve outcomes for cancer patients further.
- Significant variations and inequalities in outcomes and access to services persist.
- People in England are less likely to develop cancer than in other high-income countries in Europe but, according to the most recent data, cancer outcomes in England have generally been worse.

---

[706] *Appendix H* synthesises the most relevant findings derived from the NAO study.
[707] National Audit Office. 2010. "Delivering the Cancer Reform Strategy." London, UK.
[708] Ibid. P. 9.

- Better information is becoming available to strengthen the evidence base for cancer services [...] however, important gaps in cancer data remain.
- There is a risk that the significant progress that has been made in improving cancer services and information will stall as a result of recent changes and pressure on resources.[709]

This overview of the evaluation and audit of the Cancer Strategy shows that this has not been an integrated process. Instead, different assessment tools have been applied to obtain evidence and to verify its effectiveness. The nature of the cancer policy promotes a strong emphasis on the use of scientific evidence that feeds continuously into the decision-making process. Not only bodies such as the DoH, the NHS, PHE, and the NAO have been involved in the generation of evidence about this disease; different NGOs systematically produce research in the cancer field and are very active in the discussion of the government's effectiveness in this area.[710] This scenario reveals a highly complex policy area in which evaluation can play a predominant role in the legitimisation of the interests pursued by stakeholders. Different issues associated with cancer affect not only the perception of the effectiveness of the government's response, but also reveal different levels of notoriety to each one of these issues, which are strongly related to politics. An example of this is the Cancer Drugs Fund[711] which has received an important amount of attention due to potential cuts.[712] This displays the complexity of cancer as a policy issue in which multiple factors are interrelated and susceptible to politics. The analysis of the hidden politics of evaluation of this case is presented in Chapter 6, taking as a basis the description presented here.

[709] —. 2015b. "Progress in improving cancer services and outcomes in England." London, UK. Pp. 7-8.

[710] To provide an example, in the report made by the NAO specialists from different institutions such as the *University of Oxford*, *Imperial College*, *King's College London*, among others participated. For the inquiry made by the PAC organisations such as the Cancer Campaigning Group, Cancer Research UK and Ovarian Cancer Action provided written evidence.

[711] The Cancer Drugs Fund (CDF) 'is money the Government has set aside to pay for cancer drugs that haven't been approved by the National Institute for Health and Care Excellence (NICE) and aren't available within the NHS in England [...] The fund is worth £340 million per year.' Cancer Research UK. 2015b. "Cancer Drugs Fund." London, UK.

[712] See: McKie, R. 2015. "David Cameron's flagship Cancer Drugs Fund 'is a waste of NHS cash' " in *The Guardian*. London, UK.

4.3 Case B: The Care-Quality Commission

The Care-Quality Commission (CQC) is an executive non-departmental public body[713] in charge of regulating the provision of health and social-care services in England, also responsible for monitoring the *Mental Health Act*. Its role in the institutional landscape is to guarantee the completion of a series of standards in the services offered to the public.

In contrast to Cancer Strategy, the CQC case regards the evaluation of a public organisation; not a specific policy (see Chapter 3, Section 3.2.3). The relevance of this is that it provides a different perspective of how evaluation works in the public sector. In the case of the evaluation and auditing of the CQC, subjective judgements are more frequent and the effectiveness of the organisation is valued on the basis of different criteria. The CQC has been, since its inception, surrounded by controversial questionings about its performance and evaluation has played an active role in its evolution. This body presents an interesting case study for two main reasons. The first one is that it is the result of the merger of previous regulatory bodies, which has inexorably impacted its consolidation as a regulator and has affected its organisational learning during its almost five years of operation. As a CQC official (UK-1) highlights, these changes produce situations when...

> ... Organisations aren't able to become mature, they are not able to grow and develop their systems and processes, to develop the skills of the people that work in them and to get some consistency.

Consequently, the CQC has faced the organisational challenges associated with the establishment of a new body, e.g. the fact that 'it didn't have the budget that the three bodies would have had' (UK-2). This is an important element to highlight because it affected the performance of the CQC and, consequently, the evaluation of this organisation.

The second reason regards the strong criticisms of its performance (especially during the first years) and the excessive attention given by the media to the

---

[713] A non-departmental public body is 'a body which has a role in the processes of national government, but is not a government department or part of one, and which accordingly operates to a greater or lesser extent at arm's length from ministers.' Cabinet Office. 2013a. "Public bodies reform." London, UK.

CQC's shortcomings that have increased the rigour of the scrutiny of this organisation. As a former CQC official recalls (UK-6): 'There would be a report every single day about something CQC hadn't done right.' The overview of this case allows looking at an organisation that has been through different processes of adjustment in the pursuit of the most effective way to operate. This change of perspective will generate different findings to make the discussion about the hidden politics of evaluation in the UK more complete and robust. As in the previous case, this section is dedicated to the description of the CQC and its evaluation process.

4.3.1 General Overview of the Care-Quality Commission

The CQC began operations in 2009 after being established by the *Health and Social Care Act 2008* (Box 4.4). This new regulator[714] was the organisation resulting from the merger of three previous regulatory bodies: the Healthcare Commission, the Commission for Social Care Inspection, and the Mental Health Act Commission.[715] The main objective of the CQC is 'to protect and promote the health, safety and welfare of people who use health and social care services.'[716] The CQC carries out multiple activities 'to make sure that hospitals, care homes, dental and general practices and other care services [...] provide people with safe, effective and high-quality care, and we encourage them to make improvements.'[717] These activities are mostly directed to guarantee the users of health and social-care services that the services received meet the minimum standards set by the government.

---

[714] According to Gray and Jenkins, in the UK 'there have been concomitant developments too in performance audit and regulation with an increasing amalgamation of regulators [...] 2009 saw the merger of the Healthcare Commission, the Commission for Social Care Inspection, and Mental Health Inspectorate into the Care Quality Commission –a mega regulator!' Gray, A., and B. Jenkins. 2011. "Policy and Evaluation: Many Powers, Many Truths." in *Evaluation. Seeking Truth or Power?*, edited by P. Eliadis, J. E. Furubo, and S. Jacob. New Jersey, US: Transaction Publishers. P. 43.

[715] Under the light of the concept of post-bureaucratic organisations, this merger can be seen as a clear modification of the organisational boundaries. Elston, T. 2012. "Developments in UK executive agencies: Re-examining the 'disaggregation–reaggregation' thesis." *Public Policy and Administration* 28(1):66-89.

[716] "Health and Social Care Act."

[717] Care Quality Commission. 2013k. "Working for CQC." London, UK.

Box 4.4 Overview of the Care Quality Commission

The CQC regulates different health sectors in England including: NHS Trusts, independent healthcare, independent ambulance, dentists, adult social care and GP services. Its main regulatory activities are:
- Setting standards of quality and safety
- Registration of health and care services
- Monitoring, inspecting and regulating care services to make sure that they continue to meet the standards
- Protecting the rights of vulnerable people, including those whose rights are restricted under the Mental Health Act
- Challenging all providers, with the worst performers getting the most attention
- Making fair and authoritative judgments, supported by the best information and evidence
- Taking appropriate action if care services are failing to meet the standards.
- Carrying out in-depth investigations to look at care across the system
- Reporting on the quality of care services, publishing clear and comprehensive information, including performance ratings to help people choose care

According to its latest report, the CQC has 2,148 staff members (955 frontline inspectors) with a funding for 2012/13 of £68.1m Grant in-aid from the DoH and £92.7m annual fees from care providers.

Source: Based on Care Quality Commission. Annual report and accounts 2012/13.

In terms of governance, the CQC has a Board that is 'the senior decision-making structure [...] it provides strategic leadership to CQC and takes collective responsibility for the long- term success of the organisation.'[718] This is composed by a non-executive Chair and up to fourteen members (executive and non-executive). As an arm's-length body of the DoH, the Permanent Secretary is accountable to Parliament for the performance of the CQC.[719] In practical terms, the DoH 'agrees CQC annual business plans and monitors CQC's financial and operational performance and risks at general and strategic level through regular performance meetings.'[720] For its operation, the CQC has a Chief Executive and three inspectorates that oversee the regulation of health and social-care services (Figure 4.7).

---

[718] —. 2013c. "CQC Corporate Governance Framework." London, UK. P. 9.
[719] The DoH and the CQC have signed a framework agreement which purpose is to define 'how the Department and the CQC will work in partnership to serve patients, people who use services, their carers, the public and the taxpayer, and how both CQC and the Department discharge their accountability responsibilities effectively.'—. 2013d. "Framework Agreement between the Department of Health and Care Quality Commission." London, UK. P. 3.
[720] Ibid. P. 10.

Figure 4.7. Organisational Structure of the CQC
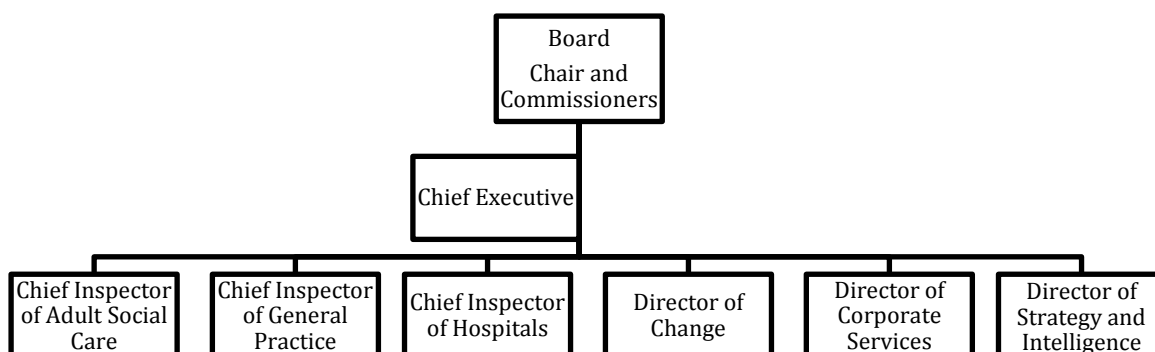
The regulatory activities of the CQC can be classified into two main processes: the registration of health and social-care providers, and the inspection of different organisations, i.e. hospitals, clinics, and dentists, to verify that the services provided meet the national standards.[721] The findings derived from these inspections are made public through the CQC's website.[722] A synthesis of the latest report of the CQC's performance can be found in Table 4.5.

Table 4.5 Key Facts of the CQC's Performance 2012/2013

Registered Providers	30,261
Inspections carried out	35,371
Whistleblowing contacts received	8,634
Warning notices to providers	910
Visits by Mental Health Act Commissioners	1,090

Source: Source: Based on Care Quality Commission. Annual report and accounts 2012/13.

This overview of the operation and governance of the CQC allows understanding more clearly its role as a regulator. Two major characteristics of this body are

---

[721] The CQC carries out two kinds of inspections: 1) *Comprehensive*, 'To make sure services are providing care that's safe, caring, effective, responsive to people's needs and well-led'; and 2) *Focused*, 'To look at something we're concerned about, which might have been raised during a comprehensive inspection or through our monitoring work [and] If there is a change in a care provider's circumstances ...' —. 2015c. "What we do on an inspection." London, UK.

722 The CQC publishes an annual report that contains information about the inspections performed during the year. In addition, it also publishes thematic reports such as *The State of Health Care and Adult Social Care in England* (October 2014).

worth highlighting: the institutional framework that aims to protect its independence for establishing judgements about the performance of health and social-care providers and the relationship between the DoH and the CQC. Regarding the latter, a CQC official stated (UK-1): 'We are independent from government and although we use taxpayers' money to discharge our function, our independence, in my view, is independence of judgement.' These two issues will become of great significance when looking at the hidden politics of evaluation in a subsequent chapter of this thesis.

4.3.2. The Evaluation of the Care-Quality Commission

The CQC is subject to the evaluation processes of its performance. This process is chaired by the Permanent Secretary of the DoH and undertaken by a panel of departmental officials and external reviewers. Its objective is 'to provide robust assurance to the public, the Department and Parliament that CQC is improving its performance and that action will be taken to build and sustain its capability in the future.'[723] This review took into consideration the main findings made by the NAO and the PAC (addressed later in this section) and it comprised the period between October 2011 and February 2012 (Table 4.6).

---

[723] Department of Health. 2012c. "Performance and Capability Review. Care Quality Commission." London, UK. P.2.

Table 4.6 Main Findings of the CQC's Performance Capability Review

1	CQC has now set the essential platform which tougher regulatory action can be taken when needed, if and where standards fall below acceptable levels.
2	The CQC has faced operational and strategic difficulties [...] delays to provider registration, shortcomings in compliance activity and, at times, a negative public profile have seriously challenged public confidence in its role.
3	There is clear evidence to show that CQC has demonstrated a new focus on its core purpose, the operational base is stronger and improving.
4	CQC must become more strategic and set out more clearly what success look like.
5	There is a blurring of the boundary between the Board and the executive team, with the Board only recently moving to take on a stronger role to constructively challenge the executive team.
6	The underlying regulatory model is new and so far there is limited practical evidence of its effectiveness.
7	The majority view from the numerous stakeholders engaged in the Review was that they want lessons to be learnt from performance shortcomings of the early years [...] recent improvements are acknowledged and there is much work to do to ensure CQC is a sustainably improving organisation delivering a respected regulatory system that protects patients and service users.
8	For the Department, the way it challenges and supports CQC also needs to be more strategic

Source: Based on Department of Health, 2012. Performance and Capability Review. Care Quality Commission. Pp. 5-6.

This instrument applied by the DoH to measure the performance of the CQC provides an example of the kind of internal evaluation that takes place. As the DoH is accountable for the CQC's performance, this review (in addition to the CQC's annual reports) constitutes evidence to be presented during external inquiries. The recommendations derived from these inquiries affect both organisations, as a member of the Health Committee staff states: 'If we say the CQC needs to rethink its purpose [...] that still has implications for the DoH and how things are regulated.' The CQC has also performed evaluations for strengthening its internal procedures. For example, it recently commissioned studies about their current regulatory model and their programme of evaluation to *The University of Manchester* and *The King's Fund*.

The CQC has also been subject to an intensive external scrutiny that has evinced a series of shortcomings in its performance and has raised an important number

of criticisms about its capacity to fit its purpose.[724] Here, it is important to highlight the role of the Health Committee as one of the bodies that recognised the need to make health regulators accountable:

> The Chair observed, when he first took on the role in 2010, that there was a gap in the accountability in the way sub regulators in the health world work [...] because they are non-departmental public bodies [...] so you have that situation where they are technically accountable to Parliament but there was no formal mechanism for Parliament to hold them to account ...(UK-2).

The identification of a gap in the accountability process for healthcare regulators is an important issue to consider because of the leverage and notoriety that the parliamentary scrutiny has had in the case of the CQC. As it is discussed in Chapter 6, the involvement of the Health Committee and the PAC has given important focus to the discussion of the CQC's performance and the need to reflect on the decisions taken by its governance body. The reviews made by the Health Committee, the PAC, and the NAO have produced a series of findings and recommendations consistent among the different scrutiny bodies (Figure 4.8).[725]

---

[724] See for example: Press Associated. 2013. "Care Quality Commission 'not fit for purpose' says BMA." in *The Guardian*. London, UK.

[725] For a more detailed description of the content of these reports, *Appendix I* summarises the main findings of each document.

Figure 4.8 Main Criticisms Made to the CQC's Performance (External Auditing)

Health Committee	•Distortion of priorities (registering providers over inspecting them) •Long-standing vacancies for CQC inspectors •Serious failings within management and organisation of the CQC •Serious failings in relation to whistleblowing
Public Accounts Committee	•DoH is responsible for the effective regulation of health and social care but has not had a grip on what the CQC has been doing •The CQC has been poorly governed and led •CQC's role is unclear •CQC must stenghten its whistleblowing arrangements
National Audit Office	•Compliace review and inspection work fell significantly during 2009-2010 and 2010-2011 •CQC underspent against its budget for 2009-10 and 2010-11, partly because it had a significant number of staff vacancies •Responsibility for funding the regulation of health and adult social care is falling increasingly on the providers of these services rather than the DoH •CQC has a systematic approach to assessing the risk that providers are not meeting the essential standards of quality and safety, but it depends on good quality information which is not always available

Source: Based on Health Committee, 2012. Accountability Hearing with the Care Quality Commission; Public Accounts Committee, 2012. The Care Quality Commission: Regulating the quality and safety of health and adult social care; and National Audit Office, 2011. The Care Quality Commission: Regulating the quality and safety of health and adult social care.

The scrutiny made of the CQC has revealed a series of failings and shortcomings in the performance of this regulator.[726] Different events related to neglected treatment to patients in hospitals and social-care homes[727] have put the CQC in the spotlight and have brought plenty of attention from the media.[728] A

---

[726] Campbell, D. 2012. "Care Quality Commission problems-timeline." in *The Guardian*. London, UK.

[727] The Winterbourne View and University Hospitals Morecambe Bay cases are relevant examples of episodes in which the approach undertaken by the CQC was highly questioned, both by Parliament, e. g. Health Committee, as well as by the press. For further reference see: BBC News. 2013. "Morecambe Bay trust failings prompts watchdog criticism." Lancashire, UK. BBC. 2011. "Undercover Care: The Abuse Exposed." London, UK.

[728] See Hitchcock, G. 2011. "Care Quality Commission faces uphill struggle, says watchdog." in *The Guardian*. London, UK.

paradigmatic case was of Winterbourne View, in which a series of abuses to patients at a private hospital were revealed by the BBC *Panorama* programme in 2011.[729] The CQC received information about these abuses from a whistle-blower and stated …

> … that the information provided was limited and different in scope to the issues raised in the programme, but recognised that, if it had contacted the whistleblower immediately, it would have been alerted to the seriousness of the situation […] instead the Commission passed the information received from the whistleblower to South Gloucestershire County Council, in November 2010, but did not follow up to check what action had been taken.[730]

Moreover, in the past, frictions between members of the CQC Board have also contributed to the politicisation of this body. The case of Kay Sheldon, a member of the Board, gained notoriety. In November 2011 she approached the Mid-Staffordshire Public Inquiry 'with serious concerns about the leadership, culture and regulatory approach of CQC.'[731] This led to a public dispute between Sheldon and the former Chair of the CQC that eventually led to the resignation of the latter in September 2012.[732]

All these issues were addressed in the external scrutiny made to the CQC. This produced the recognition of the problems that the organisation had faced since its beginning. In the words of the former Chief Executive of the CQC:

> As the NAO report makes clear, we faced a difficult task. We had to bring together the work of three organisations and bring in a new model of regulating health and adult social care. Not everything has gone smoothly, but we have learned, reviewed what we do and made changes – often with support of others involved in health and social care.[733]

An important remark is that these recommendations had an effect on the way the CQC is organised and operates at present. The publication of the document *Raising Standards, Putting People First* was an explicit attempt of the CQC to take

---

[729] Curtis, P., and H. Mulholland. Ibid."Panorama care home abuse investigation prompts government review ".
[730] National Audit Office. 2011a. "The Care Quality Commission: Regulating the quality and safety of health and adult social care." London, UK. Pp. 41-42.
[731] Health Committee. 2014a. "2013 Accountability Hearing with the Care Quality Commission." London, UK.
[732] See: BBC News. 2012. "Dame Jo Williams: Care regulator chair to resign." London, UK.
[733] Care Quality Commission. 2011a. "Care Quality Commission response to the National Audit Office Report."

into consideration the criticisms received and to establish a new strategy for the organisation. For example, one of the main criticisms was its poor performance in inspections. As a response, the CQC developed a new strategy that took into consideration these recommendations and promoted the adoption of a new scheme:

> We are planning to appoint a Chief Inspector of Hospitals, and a Chief Inspector of Social Care and Support, and considering appointing a chief inspector for primary and integrated care.[734]

These changes involved also the development of a public consultation that …

> … sets out the principles underlying how CQC will inspect all services and some more detailed proposals for how we will inspect NHS trusts and foundation trusts and independent acute hospitals […] It also includes some joint proposals between CQC and the Department of Health on changes to regulations that underpin our work.[735]

Another relevant issue derived from this scrutiny were the changes made in relation to whistle-blowers. The CQC's strategy recognised: "We will take full account of information from care staff, including 'whistleblowers'."[736] The observations made by the Health Committee also highlighted the need to revise the registration process undertaken by the CQC. This has translated in the development of a new registration process, a priority issue for the CQC, in the aim to become more efficient.[737]

More recently, the Health Committee has carried out a new hearing with the CQC to review the progress that the organisation has made over the last year. The focus of the inquiry was the VFM of the CQC and identifying whether this body was fit for purpose. Recalling the findings made in the past by the Health Committee, but also by other external reports, e.g. *The Francis Report*,[738] the

---

[734] —. 2013i. "Raising standards, putting people first. Our strategy for 2013-2016.". London, UK. P. 2.

[735] —. 2013g. "A new start. Consultation on the changes to the way the CQC regulates, inspects and monitors care." London, UK. P. 2.

[736] —. 2013i. "Raising standards, putting people first. Our strategy for 2013-2016.". London, UK. P. 12.

[737] —. 2014a. "Business Plan 2014/15 to 2015/16." London, UK.

[738] This report, published in 2013, 'examined the causes of the failings in care at Mid Staffordshire NHS Foundation Trust between 2005-2009 [it] makes 290 recommendations, including: openness, transparency and candour throughout the healthcare system (including a

core question was about the changes that the CQC has undertaken in recent times. In this sense, the current Chair of the CQC recognised that a number of important changes have taken place in the CQC, both in its governance and in the methodology to register and inspect social services. In particular, during the hearing with the Health Committee he stated:

> We have changed the entire top team in the CQC. We have changed the organisation and we have changed our governance structure. On looking outwards, we have now an intelligence-driven, risk-driven inspection regime, which has made a huge difference, assessing the risk of acute hospitals and GP practices from information from data.[739]

Another important issue refers to the incapability of the CQC to recruit inspectors. This has been a recurrent finding in the audit processes of the CQC[740] and, according to the evidence discussed in the hearing, is still an unsolved issue. Regarding this, the Chair of the CQC exposed:

> I do believe we are capable of finding them. We have tried a number of different approaches to recruitment. We have tried recruiting through a mass recruitment exercise and now we are undertaking much more bespoke recruitment exercises, recruiting locally, but also looking flexibly at different models of recruitment [...] But we have set a very high bar for inspectors. One of the criticisms we have had historically is about the quality of staff.[741]

Finally, the CQC also made explicit the relevance given to data. The development of an intelligent monitoring tool is an example of the new approach undertaken in relation to inspections. 'Our view was that there was a collection of [...] data that could be used to flag if there was a risk that the service might not meet safety and quality standards.'[742] This overview of the evaluation and audit processes of the CQC offers an interesting object of analysis of the hidden politics of evaluation because of the nature and characteristics of this

---

statutory duty of candour), fundamental standards for healthcare providers improved support for compassionate caring and committed care and stronger healthcare leadership.' The Health Foundation. 2014. "About the Francis Inquiry."

[739] Health Committee. 2014d. "2014 Accountability hearing with the Care Quality Commission." London, UK: House of Commons. P. 1.

[740] See: —. 2011a. "Accountability hearing with the Care Quality Commission." London, UK: House of Commons, —. 2013a. "2013 Annual accountability hearing with the Care Quality Commission." London, UK: House of Commons.

[741] —. 2014d. "2014 Accountability hearing with the Care Quality Commission." London, UK: House of Commons. P. 6.

[742] Ibid. P. 12.

organisation.[743] The core argument derived from this description is that the evaluation of the performance of the CQC has been more debatable because it has been based on subjective judgements. This matters for the research question of this thesis because it provides the opportunity to analyse and reflect upon how judgements are made in a context in which multiple interests are at stake and when the interpretation of success can rely on different criteria.

The next chapter describes the institutional framework for evaluation in Mexico and the case studies in this country.

---

[743] In July 2015, the NAO published a new report about the CQC's progress 'in putting its transformation strategy in place, and its capacity to implement its new approach.' The report mostly focuses on appraising the new regulatory model. National Audit Office. 2015a. "Capacity and capability to regulate the quality and safety of health and adult social care." London , UK.

Chapter 5. Case 2: Policy Evaluation in Mexico

The description made of the evaluation and audit systems in the UK (see Chapter 4) was useful for the identification of those elements of the institutional framework underpinning these activities. This chapter describes evaluation governance in Mexico to contextualise the analysis of the hidden politics of evaluation in the proceeding chapter.

This overview comprises the identification of the most relevant legal instruments for evaluation and audit at the national level. It also incorporates the description of the selected case studies: the Health-Quality System Programme (SICALIDAD) and the Health Caravans Programme.

5.1 Audit and Evaluation in Mexico

The development of evaluation and auditing in Mexico is regulated by different legal instruments that establish obligations for ministries to evaluate their policies. At the national level, is it possible to identify two major processes for evaluating and auditing public policy. The first one is an internal process implemented by the Federal Government. The second process is an external audit carried out by the Supreme Audit Office (SAO), a technical body of the Chamber of Deputies.[744] Policy evaluation in the Mexican case has been a relatively recent process, although sporadic efforts in this matter have been documented since the late seventies.[745] However, the promulgation of the *General Social Development Law* in 2004 fostered the institutionalisation of evaluation in this country. In the audit field, the establishment of the SAO as it currently operates occurred in the early 2000s after the promulgation of the *Superior Audit Federal Law*.

The following section describes in more detail the design and functioning of evaluation and auditing systems in Mexico. This overview identifies rules and procedures established for this purpose and the main stakeholders involved.

---

[744] Supreme Audit Office. 2014a. "About us."
[745] Castro, M. F., G. Lopez-Acevedo, G. Beker Busjeet, and X. Fernandez-Ordonez. 2009. "Mexico's M&E System: Scaling Up from the Sectoral to the National Level." Washington, D.C., US: The World Bank.

The Mexican case illustrates how politics affects evaluation when this activity takes place in an under-resourced, fragmented, and strongly centralised context, and how this interferes with its ultimate purpose: improving policymaking and accountability.[746]

5.1.1 The Institutional Framework for Policy Evaluation in Mexico

As stated, there are two major processes for the evaluation of public policy in Mexico: internal evaluation and external auditing. Internal evaluation is a process undertaken by the Federal Government through three coordination agencies in charge of planning, supervising, and funding evaluation. International organisations such as the World Bank and the Inter-American Development Bank have played an important role in the promotion of evaluation in Mexico, both by funding evaluation[747] as well as evaluators in certain cases.[748]

At least discursively, evaluation promoted by the Federal Government is external, as it is commissioned to evaluators outside the governmental sphere. In practice, this is questionable because the majority of the frameworks used for the evaluation of programmes are pre-established, approved, and funded by the Federal Government. Consequently, the level of influence of government impacts on the extent to which evaluation can be considered an external process. The institutionalisation of evaluation has been an incremental process. It has transited from the development of performance indicators to the establishment of specific criteria for valuing different aspects of public

---

[746] According to Mackay, a monitoring and evaluation system should serve governments for two major purposes: to support the policymaking process, particularly in the budget area, and in the national planning; and, to incentive transparency and to support accountability. Mackay, K. 2007. "How to build M & E systems to support better government." Pp. ix, 157 p. Washington, D.C., US: The World Bank.

[747] For example, they have funded evaluations and been enthusiastic sponsors of the adoption of the Results-based Management Model (RbM). The Mexican Federal Government had a loan from the World Bank for US$17.24 million for a project to foster this model in Mexico. Source: World Bank. Projects and Operations. MX Results-based Management and Budgeting. The World Bank. 2008. "MX Results-based Management and Budgeting."

[748] An example is the evaluation made by the IADB to the Mexican Rural Policy in 2007. See: Lozano, E., N. Gutiérrez, R. Díaz, and A. González. 2007. "Adquisición de bienes privados con recursos públicos. Revisión de Política Agropecuaria y Rural en México.": Banco Interamericano de Desarrollo.

programmes.[749] This has been a fragmented process observable, for example, in the health sector, in which 'evaluation areas have existed and disappeared according to the government in office' (MX-3). This might suggest that the institutionalisation of evaluation has been a process linked to the political priorities of key stakeholders, i.e. the party in office, and their intentions to incorporate this activity into the logic of the public sector. Evaluation became a formal government's policy when 'obligations to evaluate the results of social programmes were established.'[750] This is an interesting point because the implementation of formal rules in this area limited the discretion of policymakers in relation to *what*, *when*, and *how* to evaluate.

The external audit process is performed by the SAO, an autonomous body of the Chamber of Deputies that assesses 'the accomplishment of the objectives and goals established in the federal programmes [...] in order to verify their performance and the proper use of federal public resources, according to the legal framework.'[751] This body makes the Federal Government accountable through the performance of annual audits to the public expenditure executed. This is the means through which the Legislative 'reviews the way and terms in which budget is executed' (MX-12). The external nature of the SAO and its connection to the Legislative can be perceived, theoretically, as a mechanism to give this body more leverage. However, as discussed later, in practice the SAO has limited power.

Both processes (internal evaluation and external audit) are constrained by a set of legal instruments that determine the obligations that ministries have in these affairs (Table 5.1).

---

[749] Feinstein, O., and G. Hernández. 2008. "El Papel de la evaluación en México." Banco Mundial.
[750] Pérez-Jácome Friscione, D. 2012. "Presupuesto basado en resultados: Origen y aplicación en México." edited by Ministry of Finance. Mexico City. P. 3.
[751] "Federal Accountability Law." Art. 15, Fraction IV.

Table 5.1 Legal Framework for the Evaluation of Federal Programmes in Mexico

Level	Legal Instrument	Policy Evaluation Highlights
Macro Level	Political Constitution of the United Mexican States	• Policy evaluation is conceived as part of the Democratic Planning System • The Constitution gives the Executive the faculty to establish the guidelines for the design, implementation and evaluation of the national goals • It also gives evaluation faculties to the Congress, through its technical body the SAO, to 'legislate for the SAO, as well as for the performance, control and evaluation procedures to be applied to the Powers of the Union and the federal agencies' (Art. 73, fraction XXIV).
Meso Level	Federal Law of Budget and Fiscal Liability	• It mandates the obligation of evaluating systematically the federal expenditure. • It establishes the creation of the Performance Evaluation System (PES) • It acknowledges the linkage between evaluation and the budget process (Results-based Management model) • This Law defines the Ministry of Finance, the Ministry of Public Function and National Council for the Evaluation of the Social Development Policy as the agencies in charge of conducting the evaluation policy
	General Law of Social Development	• CONEVAL is the agency in charge of coordinating the evaluation of social policy
	Federal Accountability Law	• This Law regulates the activity of the SAO • It establishes the legal premises for the control of public resources and for the performance of audits
Micro Level	General Guidelines for the Evaluation of the Federal Programmes of the FPA	• This document defines the general rules for the performance of evaluations to federal programmes • It incorporates information about the administrative procedure to follow, and the nature of the evaluation approaches to be used
	Mechanism for the follow-up aspects susceptible to improvement	• It establishes the procedure for the use of recommendations derived from evaluation studies

As it is possible to observe, the legal framework is highly complex. Obligations are dispersed throughout multiple regulatory instruments and demand the participation of various ministries at different stages of the process. Despite the establishment of these formal guidelines, the legal framework presents some grey areas that have produced ambiguity about how responsibilities are allocated in terms of evaluation. In specifics, the evaluation promoted by the Federal Government is ruled by the following instruments:

- Political Constitution of the United Mexican States
- Federal Law of Budget and Fiscal Liability
- General Law of Social Development
- General Guidelines for the Evaluation of the Federal Programmes (hereafter, Guidelines for Evaluation)
- Mechanism for the follow-up aspects susceptible to improvement (hereafter, Mechanism)

In addition, the work of the SAO is regulated by the *Constitution* and the *Federal Accountability Law.*[752]

The evaluation of public programmes is based on the application of instruments that focus on values like efficiency, efficacy, and economy. These instruments are directed to obtain 'objective' evidence about the accomplishment of programmes. For example, the Specific Performance Evaluation (SPE) implies the analysis of indicators like coverage and population served to capture the level of accomplishment obtained. Evaluation has been mostly constrained at the programmatic level rather than the discussion of broader policy issues, which might generate political reactions.[753] According to a study made by GESOC, a Mexican public policy think tank, the limited scope of the evaluation policy represents one of its most notorious weaknesses:

> There are no institutional and normative grounds for performing an integrated systemic evaluation, both vertical and horizontally, bringing together the macro (policies), meso (programmes and organisations) and

---

[752] In addition to the legal framework described here, both internal and external mechanisms of evaluation may refer to other legal instruments in the development of their activities (e.g. SAO's internal regulations).

[753] For instance, in the Annual Evaluation Programme (AEP) 2015, the vast majority of the content concerns the evaluation of programmes, with the exception of the evaluation of the National Crusade against Hunger which is considered a strategic evaluation that exceeds the time framework of the AEP 2015.

micro (officials performance) levels.[754]

In this scenario, evaluation seems to be a mechanism that avoids affecting political interests at a broader level. As a result, this has restricted its effect on policymaking and accountability and it has translated it into an activity that public officials perceive as an additional administrative load that they are obliged to fulfil. According to a Ministry of Health official (MX-6) evaluation 'generates a managerial effort and this affects more those [ministries] less organised.' This resonates with the asymmetry of capabilities among actors within government and with the different conceptualisations about the utility of evaluation. The work of the SAO is more related to the control of public funds and accountability. This body audits public expenditure (after it has been spent) on an annual basis.

In both cases, the effect of these activities in policymaking and in the political debate demands a more careful description. In contrast to the UK (see Chapter 4), the Mexican institutional framework for evaluation and auditing is characterised by multiple rules and procedures that aim to guide the behaviour of the stakeholders involved in this process. These rules and institutions are described more carefully in the next sections.

5.1.2 Internal Evaluation

Internal evaluation is implemented by the Federal Government through three agencies: the National Council for the Evaluation of the Social Development Policy (CONEVAL), the Ministry of Finance, and the Ministry of Public Function. All these agencies have obligations established in the legal framework related to the planning and coordination of evaluation at the national level.

The Ministry of Finance is responsible for implementing and supervising the Performance Evaluation System (PES).[755] It also coordinates the evaluation of

---

[754] Arreola, A. (Ed.). 2008. *¿Gobernar por resultados? Implicaciones de la política de evaluación del desempeño del gobierno mexicano.* Mexico City: GESOC, A.C. P. 241.

[755] The PES is 'group of methodological elements that allows making an objective assessment of the Programmes' performance, guided by the principles of verification of the level of accomplishment of objectives and goals, based on strategic and management indicators that show the social impact of programmes and projects.' "Federal Law of Budget and Fiscal Liability."

non-social federal programmes, e.g. the communications and transport sectors. Evaluation from this ministry's perspective is used …

> … to determine the pertinence and accomplishment of objectives, the efficiency, effectiveness, impact and sustainability of these, providing credible and useful information about the obtained results to improve the decision-making process.[756]

In this context, the PES is an instrument that integrates information about public programmes and its relationship with national priorities considering performance indicators, evaluation outputs, and follow-up actions that can provide a more comprehensive overview of the accomplishments made by the government.[757] The Ministry of Public Function is also involved in the monitoring of this system. However, its tasks are more linked to auditing and control. The aim of its evaluation unit is …

> … to design and develop instruments and methodologies for an integral evaluation of the public sector performance […] that comprises the issuing and publication of guidelines, general criteria, systematic and permanent coordination mechanisms that allow establishing a public evaluation culture.[758]

Even though in 2012, after the change of government occurred, new President Enrique Peña Nieto issued a decree that eliminates the Ministry of Public Function, and its functions and responsibilities were presumably transferred to the Secretariat of the Interior. Until this day its situation is unclear within the Federal Public Administration (FPA).[759] Apparently, it continues operating as in previous years, e.g. according to the current Annual Evaluation Programme (AEP), the Ministry of Public Function is in charge of coordinating the evaluation of 37 federal programmes, including some implemented by the Ministry of Finance.

The third agency involved in the execution of the Mexican evaluation policy is CONEVAL. This is 'a public decentralized organisation of the Federal Public Administration with legal personality and own equity, with technical and

---

[756] Ministry of Finance. 2011. "Results-based Budget. Evaluation.".
[757] —. "Cédula: Conceptos Básicos del Presupuesto Basado en Resultados. Sistema de Evaluación del Desempeño (PbR-SED)."
[758] Ministry of Public Function. 2014. "Government Performance Evaluation Unit." Mexico City.
[759] "Decree that amends, supplements or repeals certain provisions of the Organic Law of the Federal Public Administration."

managerial autonomy [...], it is within the sector coordinated by the Ministry of Social Development'.[760] Its role is evaluating social policy at the national level and measuring systematically poverty indicators in the country. Its establishment was the result of an arrangement between political parties that led to the publication of the *General Social Development Law* in 2004.[761]

The institutional design establishes these three actors as the organisations responsible for conducting the evaluation policy. In empirical terms, this set of arrangements has produced conflicts between these agencies and has affected the relationship with the evaluated bodies. According to a Ministry of Public Function official (MX-2), this design 'is totally inefficient because you have three structures in three different institutions with three different approaches'. Moreover, the legal framework does not set formal boundaries for the separation of the competencies of each agency. The legal framework, for example, limits CONEVAL's scope to social policy. However, this concept is not clearly defined in the law and it has generated ambiguity among coordination agencies.[762] In addition, as a CONEVAL official recognises (MX-9) this has also produced duplicities. 'We have performed very similar functions, with very similar objectives and I think it is a consequence of the fact that it has not defined very clearly the scope of each agency.'

Furthermore, the internal evaluation process raises important concerns about issues like independence, impartiality, and credibility, addressed in more detail in Chapter 6. Yet, at this point it is important to acknowledge that the process implies the interaction of these three agencies with the evaluated object (an

---

[760] CONEVAL is a decentralized body that can be compared to quangos in the UK. "Decree that regulates the National Council for the Evaluation of the Social Development Policy." CONEVAL is a decentralized body that can be compared to quangos in the UK.

[761] Hernández Licona, G. 2010. "Construyendo un Sistema de Evaluación y Medición: Un Reto de Política Pública con Elementos Técnicos."

[762] Due to the gaps found in the legal framework, the AEP 2011 presents definitions for 'federal actions of social development' and 'federal programme of social development' with the purpose of making clearer the influence of CONEVAL and its scope in the evaluation policy. In addition, in 2011 CONEVAL issued an *Inventory of Federal Social Actions and Programmes* which aim was to integrate the information about this kind of programmes implemented by the Federal Government. This inventory has registered 273 social programmes or actions. In practical terms, these three agencies have distributed the universe of federal programmes based on the classification used for budget purposes and for public expenditure control to carry out evaluation activities.

operative unit of the programme) and with an evaluation unit within each public agency whose main functions include 'hiring, operating and supervising evaluation.'[763] Internal evaluation relies strongly on the capabilities (technical and political) of different stakeholders within the Federal Government. Evaluation planning is fundamentally concentrated in the coordination agencies, while the implementation and the use of evaluation outputs are under the scope of the evaluated agencies.

The lack of boundaries in the evaluation policy has generated two major problems. First, it has created power struggles that can be observed in the discretionary decisions made during the process—negotiation for the evaluation space. For example, the selection of the programmes to evaluate during the fiscal year appears to be more a negotiation between these parties than a decision based on explicit and transparent criteria. There is no explicit criterion for including or dismissing programmes, neither to decide which will be the coordination agency in charge. According to a member of the Ministry of Finance (MX-19): 'Anyone eliminates or adds those that he likes.' For instance, the AEP2015 establishes the performance of a cost-effectiveness evaluation coordinated by the Ministry of Finance for the Social Milk Supply Programme (LICONSA, its acronym in Spanish), while it also sets the revision of the programme's indicator matrix and a design evaluation coordinated by CONEVAL. This exposes the lack of clear criteria for the selection of programmes to evaluate. LICONSA is considered a social programme; under the scope of CONEVAL, although the Ministry of Finance is also involved in its evaluation.

This selection is left as a discretionary decision of these three agencies and it can leave room for using evaluation as a political tool to evince the weaknesses or the strengths of a programme. To illustrate this, the Human Development Programme *Oportunidades* has been considered in every AEP since 2007. This was the most iconic programme of President Calderón's Government (2006–12)

---

[763] The Guidelines for Evaluation establish that public agencies within the FPA, 'without modifying their organisational structure, must designate an area [...] that will be responsible for supervising the quality and the normative enforcement of evaluations, as well as for the coordination with CONEVAL ...' *Guidelines for evaluation* (n. 29).

and there was a clear interest in promoting its achievement. The international recognition obtained by the programme was also an important factor for its legitimisation.[764] International organisations play an important role in this— 'They [programmes] bring in the World Bank to perform their evaluation [...] that means that they do not believe in the evaluation that takes place here' (MX-2). This reveals that the value given to the credibility and legitimacy of evaluation differs in relation to the organisation performing the study. From this perspective, the involvement of an international organisation may provide more leverage to the findings of a study.

The second consequence of this discretion is that there are incentives for evaluating policy areas in which more evaluation capabilities have been built, e.g. social policy (MX-9). Just to illustrate, the AEP2013 established the evaluation of 202 federal programmes during this fiscal year; only 24% are non-social. In addition, social programmes are considered more politically profitable, hence it is important to promote and disseminate achievements in this area. Again, the cases of the *Oportunidades* and *Seguro Popular* programmes are paradigmatic, as they have an explicit evaluation agenda, in contrast to the rest of the federal programmes operating.[765] The evaluation agenda of these programmes transcends the identification of goals as it also involves the development of more specific studies to observe their effect on vulnerable sectors of the population, i.e. indigenous people.[766] The political legitimisation of policies based on evaluation evidence is a common practice, particularly in those contexts in which the programme to be legitimised represents an

[764] See: The World Bank. 2014. "A Model from Mexico to the World." .

[765] See: Coordinación Nacional del Programa de Desarrollo Humano Oportunidades. 2007. "Agenda Integral de Evaluación del Programa Oportunidades 2007 - 2012 (Evaluation agenda for the *Oportunidades* Programme 2007-2012)." Mexico City: SEDESOL, Ministry of Health. 2005. "Sistema de Protección Social en Salud. Estrategia de Evaluación (Evaluation Strategy)." Mexico City.

[766] See: Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS). 2010. "Indigenous Communication Programme Evaluation." Mexico City. and Sosa-Rubí, S., A. Salinas-Rodríguez, and O. Galárraga. 2011. "Impacto del Seguro Popular en el gasto catastrófico y de bolsillo en el México rural y urbano, 2005-2008 (Impact on the catastrophic expenditure of the *Seguro Popular* programme in the rural and urban Mexico, 2005-2008)." Mexico City: Instituto Nacional de Salud Pública.

important amount of public funds.[767]

In spite of the ambiguity and complexity of the legal framework, it defines the general parameters of the administrative process to follow for the evaluation of federal programmes. The *Budget Law,* for instance, establishes some important issues worth mentioning:

- Evaluation to federal programmes can be performed either by public agencies itself or by specialised external agents who meet the requirements of independence, impartiality, transparency and any others established in the legal framework
- Evaluations will be public and must provide information about the general characteristics of the evaluation to perform—information about the external evaluator team, general data of the administrative unit within the public agency in charge of coordinating and monitoring the evaluation, the hiring scheme of the evaluation, databases, data-collection instruments, e.g. questionnaires or interviews, methodology, and the main characteristics of the sample, executive summary containing the main results obtained, recommendations from the external evaluator, and the total cost of the evaluation with its source of funding
- Evaluation can be performed at different levels (policies, programmes, and institutions)
- It is mandatory to establish an *AEP*
- Federal programmes must follow recommendations derived from evaluations.[768]

In addition, the *Social Development Law* defines CONEVAL as the agency in charge of evaluating of social policy[769] by 'revising periodically the accomplishment of the social objective of the programmes, goals and actions of the social policy with the purpose of correcting, modifying, adding, reorienting or supressing them totally or partially' (Art. 72).[770]

---

[767] According to a study made by GESOC, *Oportunidades* and *Seguro Popular* were the programmes with the largest budget in the Federal Expenditure Budget (circa 3 billion GBP each). Source: GESOC, A.C. 2014. "Índice de Desempeño de Programas Públicos Federales 2014 (Performance Index of Federal Programmes 2014)." Mexico City.

[768] "Federal Law of Budget and Fiscal Liability." Art. 110.

[769] The Social Development Law (2004) acknowledges CONEVAL as the agency in charge of the evaluation of social policy; however, as a former CONEVAL's counsellor has identified, the Budget Law limited the scope of CONEVAL by giving joint evaluation faculties to the Ministry of Finance and the Ministry of Public Function, reducing its capacity of influence and its power. Acosta, F. 2010. "La evaluación de la política social en México: Avances recientes, tareas pendientes y dilemas persistentes." *Papeles de Población* 64:155-88.

[770] Social policy represents the most visible policy field in which evaluation has been undertaken, where more assessment instruments have been applied, and with more visible results.

- Higher education institutions—scientific or non-profit organisations—can evaluate social policy, CONEVAL can perform evaluations by itself, but when an external organisation is needed for the evaluation CONEVAL must do a public bidding
- Evaluation must be annual; it can be multiannual when it is considered necessary
- Evaluation outputs must incorporate the results, management, and service indicators to assess the coverage, quality, and impact of programmes
- Evaluation outputs must be published in the Official Federation Journal, as well as sent to the Social Development Commissions of the Chamber of Deputies and Senators
- CONEVAL, based on the results of evaluations, can make recommendations to the Executive and make them public.[771]

Box 5.1 General Overview of CONEVAL

Through a Presidential Decree issued in 2005, CONEVAL was created as public decentralized organisation of the Federal Public Administration with two main purposes:
1. To regulate and coordinate the evaluation of the social development policy.
2. To establish the guidelines and criteria for the definition, identification and measuring of poverty.

According to the WB, CONEVAL 'is the lynchpin of the social sector M&E system in Mexico […] it coordinates cross-sectorial evaluation activities, sets out standards and methodologies for the evaluation of social Programmes, provides technical assistance to ministries and agencies and undertakes or manages some specific evaluations'.

CONEVAL's Secretary is named by the Federal Government; however, the governance of this body relies on a Directive Committee composed by the Secretary of the Ministry of Social Development, six academic researchers, a representative of the Ministry of Finance, a technical secretary and a prosecretary. The Executive Commission, composed by the Executive Secretary and six academic researchers, which are chosen through a public process, is the technical body that guides the operation of CONEVAL. It has an administrative structure of 65 public officers distributed into five administrative areas: Executive Secretariat, General Direction of Coordination, General Direction of Administration, General Direction of Evaluation and General Direction of Poverty Measurement. According to information of the Ministry of Finance, for the 2015 fiscal year, CONEVAL has a budget of 440 million pesos (approximately £18 million) and it will coordinate the performance of 158 evaluations according to the AEP2015 (45 designs evaluations and 113 performance evaluations).

Source: Based on CONEVAL and Castro et al., 2009 *. Mexico's M&E System: Scaling Up from the Sectoral to the National Level.* The Wolrd Bank.

As a consequence of the complexity of the legal framework,[772] coordination agencies published in 2007 the *Guidelines for Evaluation* to standardise the evaluation process (Table 5.2). Before this institutionalisation process, as a former CONEVAL's official explains (MX-24), information about the

---

[771] "Social Development General Law." Mexico. Title Fifth, Chapter I.

[772] In addition to the legal framework described in this section, there is a set of instruments that aim to define the processes for the development of evaluation to federal programmes and to 'regulate' the work of evaluators by establishing *ex ante* the expected outputs of the evaluation, the obligations of the parties involved. These instruments are commonly known as Terms of Reference (ToR) and are usually developed by the unit in charge of coordinating the evaluation.

effectiveness of programmes was dispersed and fragmented and 'CONEVAL needed to generate instruments for this purpose'. This underlines the leverage given to its role as the agency responsible for promoting the use of evaluation, but also for generating evaluation capabilities within the public sector. CONEVAL lacks the capacity for enforcing its recommendations because unlike its counterparts in the evaluation system (Ministries of Public Function and of Finance) it is not associated with a sanction component. Instead, as a CONEVAL official (MX-9) recognises, 'We have legitimacy […] we are academics […] this backs us up'. CONEVAL's governance, as discussed in Chapter 6, has influenced the way in which this body is perceived. Regardless of sitting within the public sector as a parastatal body, it has built an important reputation as an institution with 'technical quality and academic independence' (MX-14).

In this context, the *Guidelines for Evaluation* provided a more structured approach to the administrative process of evaluation and its nature. Here, evaluation is defined as 'the systematic and objective analysis of federal programmes whose purpose is to determine its pertinence and the accomplishment of its objectives, as well as its efficiency, efficacy, quality, results, impact and sustainability' (Table 5.2).[773]

---

[773] "General guidelines for the evaluation of the federal programmes of the Federal Public Administration." Mexico.

Table 5.2 General Content of the *Guidelines for Evaluation*

Monitoring	Evaluation Approaches	Recommendations derived from evaluation	Dissemination of evaluation results	Federal Programmes Integral Monitoring and Evaluation System

All federal programmes must develop an indicators matrix that will be incorporated into the Performance Evaluation System.
 	
The methodology used for this matrix will be the logical framework approach.
	The following evaluation approaches are addressed in the Guidelines for Evaluation: Design Evaluation (mandatory for those new programmes during the first year of operation). Consistency and Results Evaluation Indicators Evaluation Processes Evaluation Impact Evaluation Specific Evaluation Strategic Evaluation Complementary Evaluation (all those evaluation not considered in the AEP).	Public agencies must heed recommendations derived from evaluations according to the model that will be established for that purpose by the Ministries of Finance, Public Function and CONEVAL.	Public agencies must disseminate the results of the through their websites.
			    They are also obligated to inform about the evaluator that was hired, the cost of the evaluation, among other aspects.
				The three parties involved in the federal evaluation policy (Ministry of Finance, Ministry of Public Function and CONEVAL) must integrate a system based on information derived from the monitoring and evaluation processes in order to provide public agencies with evidence about the use of public resources, the advancements and achievements made in their results.
				    This system must contribute to the programme's improvement, to the decision-making process and to the accountability process.

Source: Guidelines for Evaluation, 2007.

212

The *Guidelines for Evaluation* consider two kinds of evaluations. The first type are those established by the AEP (mandatory) that specify the kind of evaluation to be performed, the coordination agency, and timescales. The second group are all those complementary/*ad hoc* evaluations performed by public agencies according to their specific needs of information. The main difference between them is basically the stakeholder that defines the guidelines for the evaluation. While for those mandatory evaluations coordination agencies define the scope and instruments to apply, the *ad hoc* evaluations are planned and designed within the evaluated agency (Figures 5.1 and 5.2).[774]

Figure 5.1 Process Followed for an Evaluation Established in the AEP

Publication of the Annual Evaluation Programme

Administrative process to hire the external evaluator

The contract is commissioned and evaluation begins based on the TOR established

Coordinator mediates the interaction between the external evaluator and the evaluated Programme

The evaluation is completed and delivered to the coordinator and the evaluated agency, and it becomes public

Evaluated programmes are obligated to follow-up recommendations according to the procedure established by the *Mechanism*

Source: Based on Guidelines for Evaluation.

---

[774] This description is a simplification of these processes only for illustration purposes. In practice, these processes are much more complex and composed by many other sub processes that are not shown here. The hiring of external evaluators, for example, is carried out based on what is established by the *Law of Acquisitions, Leases, and Services of the Public Sector*, so this constitutes a process that involves a series of administrative activities additional to what it is shown here.

Figure 5.2 Process Followed for a Complementary (*ad hoc*) Evaluation

Public agencies decide to perform an evaluation to their programmes based on the needs identified and the availability of resources for this purpose

The evaluation unit develops the TOR(CONEVAL and Ministries of Finance and of Public Function might express an opinion) and carries out the hiring process for the evaluator

The evaluation is assigned to an external evaluator and it begins based on the TOR established

The evaluation unit mediates the interaction between the external evaluator and the operation unit of the programme

The evaluation is completed and delivered to evaluation unit and it becomes public

Evaluated Programmes are obligated to follow-up recommendations according to the procedure established by the *Mechanism*

Source: Based on Guidelines for Evaluation.

The process defined by the legal framework for the evaluation of federal programmes is complemented by the *Mechanism for the Follow-up of Aspects Susceptible to Improvement*, the primary objectives of which are:

> To establish the procedure for the follow-up of recommendations by external evaluations to make an effective use of evaluation finds in the improvement of programmes and policies [...] to define the persons responsible for establishing the work instruments to provide follow-up of the aspects susceptible to improvement, as well as for their execution; and to define the deadlines for the follow-up of the aspects susceptible to improvement documents[775].

This *Mechanism* outlines a procedure based on the analysis and classification of recommendations from evaluation reports. The evaluated programme analyses recommendations and classifies them according to the actors involved in their attention, e.g. operation units, other government agencies, or state or municipal governments, and according to its level of priority (high, medium, or low).[776] Through this procedure, ministries define the actions to implement, the responsibility of those actions, and deadlines for the accomplishment. This is the only formal instrument of the Federal Government to foster the use of evaluations in policymaking. However, it lacks enforcement capacity, and it has

---

[775] "Mechanism for the Follow-up of Aspects Susceptible to Improvement." N. 1.
[776] This also implies the generation of a set of documents in which public entities define the actions that they are committed to perform, as well as reporting on an annual basis to CONEVAL and the Ministry of Public Function the accomplishment of the actions that they established.

produced limited results because it is subject to the discretion of operators. As an official of the Ministry of Health (MX-4) recognises, they 'need to be very careful in the selection of the issues that they will work on […] some cheat a little, but generally speaking they choose those things they can accomplish'. The effectiveness of the *Mechanism* relies on the operator's decisions about which recommendations to follow. As the compliance of this process is overviewed by coordination agencies, operators have incentives 'not to choose the hardest ones' (MX-8). Some aspects require only rephrasing or rethinking how they report information about the programme. Consequently, they can be easily accomplished and reported (see Box 5.2).

Box 5.2 Example of the Follow-up of Aspects Susceptible to Improvement

After the delivery of an evaluation report, the evaluated programme, along with other actors considered relevant for this purpose, analyse and discuss the findings and recommendations made by the evaluator. As a consequence of this process, the evaluate issues the following documents:

- An institutional statement on the findings made by the evaluator
- A work plan to follow-up the aspects considered direct responsibility of the operative unit of the programme (known as specific aspects)
- An institutional work plan to follow-up the aspects considered responsibility of different areas of the Ministry (known as institutional aspects)

The evaluand must inform about the progress in the attention of the recommendations, as it can be observed in this example of the Health Caravans Programme:

Aspect susceptible to improvement: "To improve the definition of the *purpose* in the Logical Framework Matrix".
Activities to perform: To modify the narrative summary of the matrix in the *purpose* component.
Responsible Area: Adjunct General Direction of Health Caravans
Due date: 31/08/2013
Expected Results: Improvement of the *purpose* component.
% Progress: 50

Source: Based on Ministry of Health, 2013. Follow-up mechanism of recommendations. Work plan. Health Caravans Programme.

This description of the institutional framework for evaluation leads to reflecting upon two main points. The first one is the centralisation of the evaluation policy and the control exercised by the Federal Government in the planning and execution of evaluations. The second concerns the level of coordination that the implementation of the evaluation policy demands. The coexistence of three agencies involved in the evaluation policy has inevitably led to confrontation in terms of asymmetries of power and resources, worsened by the imprecisions of

the institutional framework about their role. In line with this, a Ministry of Health official (MX5) adds: 'They don't reach agreements […] they want to state who owns the evaluation policy and who gets to say how is it done.' These power disputes are derived from the fact that each stakeholder has an agenda to pursue. This is discussed carefully in the analytical chapter. The next section focuses on the process for the external auditing of public policy in Mexico.

5.1.3 External Auditing

The external auditing of federal programmes is undertaken by the SAO, a technical organisation linked to the Chamber of Deputies. According to the *Accountability Law* performance auditing is 'the verification of the accomplishment of the objectives established in federal programmes through the estimation or calculation of the obtained results in qualitative or quantitative terms, or both.'[777] The nature of this law and of the SAO itself is more related to control of the use of public resources rather than to evaluation. However, recently audit organisations have taken performance assessment as a new dimension of their work, leading to a debate about the differences between evaluation and auditing (see Chapter 1, Section 1.1). In this sense, the SAO has the authority for 'evaluating the final accomplishment of the objectives and goals of federal programmes […] with the purpose of verifying its performance and, if it is the case, the use of public resources according to the legal framework, independently of the fact that other instances might have similar attributions' (Art. 15 fraction IV). As an external body, the SAO verifies the appropriate use of public funds and oversees that public organisations comply with the tasks established by the law. This body was conceived as a 'counterpart to the Executive' (MX-12).

---

[777] "Federal Accountability Law." Art. 2.

Box 5.3 General Overview of the SAO

The SAO is the audit organisation in which the Chamber of Deputies deposits the faculty of verifying the Federal Public Account at the end of every fiscal year. Some of its most relevant faculties are:

- To perform audits to public institutions funded by federal resources
- To establish fines and sanctions
- To publish its reports
- To make public institutions accountable
- To establish rules and procedures, methods and accountability systems to control public income and expenditure

The Chamber of Deputies selects the head of the SAO, the Supreme Auditor remains in office for 8 years, with the possibility to be elected for one extra period. The Supreme Audit Office Surveillance Commission of the Chamber of Deputies is the body in charge of supervising the work of the SAO. According to the Federal Budget Decree, for the 2015 fiscal year, the SAO has a budget of approximately £87 million. There are about 1,500 people working for the SAO, in addition to external consultants hired for different projects. The most recent report issued by the SAO for the Federal Public Account 2013 establishes that for this fiscal year the SAO carried out 877 financial audits, 525 performance audits, 8 evaluations and 3 studies.

Source: Based on SAO. *About us.*

Overall, the SAO's audits aim to assess multiple features of public policy. These include the performance of financial audits, public work audits, and follow-up audits. Performance audit is the most similar instrument to what is done by the Executive in terms of evaluation; it focuses on the analysis of …

> … the accomplishment of the audited entity of the objectives and goals set in the evaluated programmes and projects; the use of the resources given for that purpose in terms of efficacy, efficiency and economy; and the accomplishment of the indicators published and approved in the Federal Expenditure Budget.[778]

The process that follows the SAO for the performance of audits depends on the nature of the analysis. However, some general stages can be identified (Figure 5.3). Its work is driven by the logic of verifying those programmes or policies considered relevant, according to the amount of federal resources that they represent, as well as its linkage to the governmental priorities established in the National Development Plan. These audits mostly highlight omissions in the observance of the law, but lack the enforcement capacity to make public officials amend these omissions. The most relevant part of the process is the way in which the SAO ensures the follow-up of their recommendations.[779] This

---

[778] Merino, M. 2009. "Informe sobre la evolución y el desempeño de la Auditoría Superior de la Federación." Mexico City. P. 21.

[779] Just like in the case of the *Mechanism* promoted by the Federal Government, the SAO also lacks instruments of enforcement for the attention of the recommendations derived from audits. The lack of legal capabilities of the SAO to enforce audited entities to take into consideration their recommendations is an issue of concern about the impact that the SAO has on the

basically consists of informing audited entities about the remarks made by the SAO and the actions they are expected to implement.

Figure 5.3 General Process Undertaken by the SAO for the Audit of the Federal Account

| Audit Planning/ Identification of audit entities | → | Audit Development | → | Report of Results | → | Follow-up of recommendations |
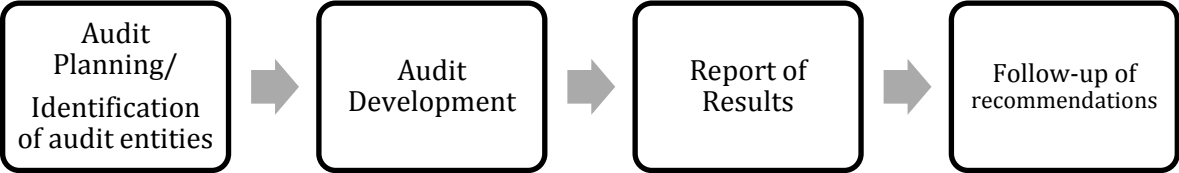
Source: Based on Merino, M., 2009. Informe sobre la evolución y el desempeño de la Auditoría Superior de la Federación.

As it occurs in the case of internal evaluation, the nature of the recommendations— especially those related to performance—concern programmatic issues that audited entities can easily implement, but that do not represent an important change in the rationale of the programme. The recommendations made by the SAO do not seem to generate significant political disturbance and their effect on the improvement of policymaking and accountability has been largely questioned. In the opinion of a Legislative researcher (MX-13), recommendations 'are sometimes driven by a logic of meeting quotas', which suggests that the relevance of the SAO's work might be constrained to perform a bureaucratic task without reflecting on its value as a means for making the Executive accountable.

This relates to the reactions about the work of the SAO produced by the recent *General Report of the Public Account 2013*. According to *El Financiero*, a national newspaper, a group of deputies (members of the opposition) strongly criticised the work of the SAO by considering it an expensive and bureaucratic organisation unable to make the government accountable and to sanction its poor performance.[780] This reanimated the debate about the pertinence of the SAO's institutional design. The audits performed by this organisation are based

---

improvement of policymaking and on accountability. As a study about the performance of the SAO suggests 'it requires more legal capacity to impose direct administrative sanctions to those public offices who have failed in their responsibilities, also the SAO needs more capacity for verifying that the changes that have been suggested are being taken into account.' ibid. P. 62.

[780] Chávez, V. 2015. "En 14 años, la ASF sólo ha castigado a siete personas (In 14 years, the SAO has only punished seven persons)." in *El Financiero*. Mexico City.

on two ideas: ex-post and yearly analysis. This has raised questions about the effectiveness of this model, as sometimes the intervention of the SAO is perceived as belated in relation to the dynamics of the public sector. An SAO auditor (MX-12) recognises that this design produces a limited impact of performance auditing because 'the law only obliges the audited body to tell us the extent to which they follow our recommendations and that will be the end of our intervention.' A deputy of the Party of the Democratic Revolution (PRD, its acronym in Spanish) agrees:

> The report on the Public Account 2013 [...] makes evident that it is requires an institutional redesign of this organisation, it is necessary to give them 'teeth' to sanction with rigor those who make inappropriate use of public funds and for them to deliver audit reports in real time in order to be able to make the amendments opportunely.[781]

In its latest report the SAO made 6,902 recommendations derived from auditing the Federal Public Account.[782] The key issue to discuss is not the number of recommendations issued and followed up, but its nature and its effect on the policymaking process and on accountability. It seems that the external audit process (at least in the performance dimension) has not permeated the broader political discussion because it does not evidence important flaws. Instead it shows minor administrative omissions that can be easily fixed by the audited entities.[783]

This has fostered the debate about the need of giving the SAO more faculties and capabilities to become an organisation that can effectively hold the government accountable. This is particularly so in the Mexican context in which corruption is considered a major problem. According to *Mexican Transparency,* 200 million acts of corruption and circa £1.4 million were used for bribing public officials at all three levels of government in Mexico.[784]

---

[781] Alcántara, S., and H. Jiménez. 2015. "Tibia, actuación de la ASF ante irregularidades: Raya (Lukewarm, SAO performance before inconsistencies: Raya)." in *El Universal*. Mexico City.
[782] Supreme Auditor. 2015. "Delivery of the General Report of the Audit to the Federal Audit Account 2013." Mexico City.
[783] See: Pelaez, M. 2015. "Aureoles pide que la ASF tenga dientes para sancionar (Aureoles asks for "teeth" for the SAO to sanction)." Azteca Noticias.
[784] Transparencia Mexicana. 2011. "Índice Nacional de Corrupción y Buen Gobierno (National Corruption and Good Government Index) Executive Summary 2010.". Mexico City.

Based on the description of the institutional framework for evaluation and auditing, it is possible to make the following remarks about the Mexican case:

- In both cases, the institutional framework is highly complex. Obligations in evaluation matters are dispersed in several legal instruments and multiple actors are involved in the implementation of the evaluation policy. In the audit field, the wide scope of policies and programmes subject to the work of the SAO makes it difficult to reach a more specific level of analysis.
- Regarding internal evaluation, the Federal Government exercises strong political power over the instruments, methodologies, and use of evaluation outputs. A group of academics and consultants dominate the market of evaluation and there are no clear standards to assess its quality.
- The scope of the analysis performed by the SAO is mostly constrained by the fulfilment of legal requirements and it does not appear to have an important effect in the improvement of programmes and policies. As an ex post analysis, the pertinence of its recommendations seems unsuitable for the dynamics of the public sector.
- Neither the bodies in charge of internal evaluation nor of external auditing seem to have the institutional capacity to enforce the use of recommendations.

Despite evaluation in Mexico being a relatively recent process, mostly constrained to the programmatic level, there is evidence that the politics has affected this activity, mostly derived from the control that the Federal Government holds and the lack of institutional strength of the SAO, which raises questions about independence and objectivity. For a clearer perspective of the implementation of the evaluation policy in Mexico, the next two sections describe two case studies that will provide empirical evidence about the hidden politics of evaluation. This overview illustrates how evaluation and audit processes for federal programmes are undertaken. The description of these cases has been made through the documentary revision of evaluation reports and government planning documents. Based on this description, it is possible to observe how the fragmentation of the evaluation policy and the relegation of relevant stakeholders have generated discretionary spaces available for political use.

5.2 Case C: The Health-Quality System Programme

This section focuses on two federal programmes taken as case studies for obtaining empirical evidence about the interaction between politics and evaluation. Both selected programmes (the Health-Quality System and the Health Caravans Programme) are part of the health policy undertaken by the Mexican Federal Government and operated by the Ministry of Health.

5.2.1 General Overview of the Health-Quality System Programme

The Health-Quality System Programme (hereafter, SICALIDAD) emerged from an initiative established in the Sectorial Health Programme 2007–12 for 'providing health services to all the population and to improve its quality by addressing priority sanitary problems and factors that determine and affect health.'[785] Its main objective is 'to strengthen the attention quality in public institutions of the National Health System through fostering the execution of projects that contribute to the effective access to health services.'[786] SICALIDAD is composed by a series of projects that address different dimensions of health quality (Table 5.3).[787]

Table 5.3 Dimensions of the Health Quality System Programme

Dimensions of Health Quality	Description
Perceived quality	It considers the citizen's expectations by implementing actions directed to improve the quality of health services perceived by citizens. This dimension also incorporates the perception of health professionals.
Technical quality and patient safety	It aims to improve technical quality and patient's safety by continuous improvement, the reduction of adverse events and evidence-based medicine.
Quality in the provision of health services/ Institutionalization of quality	To conduct, with the leadership of the Ministry of Health, the changes in the quality of the provision of health services, including the accreditation system, quality guarantee and organizational improvement.

Source: Based on Ministry of Health. SICALIDAD Specific Action Programme 2007-2012.

---

[785] Ministry of Health. 2007a. "Sectorial Health Programme 2007-2012 ", edited by Ministry of Health. Mexico City.

[786] "Agreement that issues the operational rules of the Integral Quality Health System for the 2014 fiscal year."

[787] The number of projects changed after the current government took office in 2012; however, in general terms the nature of these projects has focused on the improvement in the quality of the services provided.

The programme provides funding to public health institutions to promote the development of health-quality improvement projects. In order to obtain SICALIDAD's support, these projects need to be consistent with the national strategies for the improvement of health quality, e.g. initiatives directed to make rational use of medicines and the implementation of the Total Quality Management Model. For its operation, SICALIDAD requires the interaction of multiple actors from different sectors and demands a high level of coordination which turns it into a highly complex programme (Figure 5.4). [788]

Figure 5.4 Overview of the Operation of SICALIDAD

| Health institutions present their projects | SICALIDAD verifies that projects fulfill the requirements established and a Comitee decides whether they are accepted or not; the decision is notified to the participants | Resources are transferred to the institutions which project was accepted and a colaboration agreement is signed | Institutions use the resources for their project and inform the Ministry of Health about the execution |

Source: Based on 2014 ORP SICALIDAD.

Considering the activities undertaken by SICALIDAD, the programme has established a series of indicators to monitor the accomplishment of its results which are reported to coordination agencies (Table 5.4).

---

[788] The operation of SICALIDAD involves the participation of multiple actors from the federal and local governments and other non-governmental actors. Each project supported has its own process of operation; the process presented here aims only to provide a general overview of the operation of the programme. For a deeper understanding of this process it is advisable to refer to the Operational Rules of the programme.

Table 5.4 Indicators of SICALIDAD

Name	Periodicity
Percentage of medical units associated to a SICALIDAD project	Annual
Percentage of satisfied users with the quality of the health services of the National Health System	Annual
Percentage of states that associate at least one hospital to the clinical integrated record project	Biannual
Percentage of certified health units that serve the Social Protection System	Annual
Quality percentage of training projects that participate in funding calls	Annual

Source: Based on 2012/2013 SPE SICALIDAD.

These indicators make explicit those elements of the programme considered as primordial to measure its objectives and goals. The next section presents an overview of the evaluation and audit processes that have been applied to SICALIDAD. This information contributes to the understanding of the logic of the evaluation policy applied to a specific programme.

5.2.2 The Evaluation of the Health-Quality System Programme

SICALIDAD is subject to the legal framework that enforces its systematic evaluation, as well as to inform periodically about how it has spent its resources. This programme has been evaluated based on the parameters established within the Federal Government (internal evaluation) (Table 5.5), but it was also audited by the SAO in 2011.

Table 5.5 Evaluations Performed on SICALIDAD 2007-2012

Year	Evaluation Type	Evaluator	Budget[789]	General Description of the Study
2007	Design Evaluation (prior operation)	National Institute of Public Health	Not Available	It assesses the consistency on the design of the Programme through a standardized framework developed by CONEVAL.
2008	Specific Performance Evaluation	Demography and Health Research, S.C.	£3,900	It is an annual synthetic review of the performance of the programme in terms of a series of management and result indicators, as well as its budget performance.
2009-2010	Specific Performance Evaluation	National Institute of Public Health	£3,100	It aims to identify key information about the performance of the programme through the analysis of its indicators. It is based on a standardized framework developed by CONEVAL.
2010-2011	Specific Performance Evaluation	Dr. Antonio Carlos Martín del Campo Rodríguez (Independent consultant)	£3,000	
2009	Complementary Evaluation (Technical, Financial and Processes Evaluation)	Metropolitan Autonomous University	£175,000	It was a technical-financial appraisal and a process evaluation. A comparative analysis of a sample projects both funded and non-funded by the programme using a mixed approach.
2010	Complementary Evaluation	National Autonomous University of Mexico	£190,000	It was an evaluation of a sample of projects in 95 hospitals and using a mixed approach based on questionnaires and interviews.
2011	Complementary Evaluation	Metropolitan Autonomous University	£107,000	It was an analysis of the positioning of quality attention as a priority of the National Health System. It was made through documentary analysis and interviews.
2012	Complementary Evaluation	Metropolitan Autonomous University	£69,000	It was a meta-evaluation of the programme, considering the evaluation reports delivered between 2008 and 2011.

Source: Based on data from the Ministry of Health.

The evaluation of SICALIDAD has basically focused on the performance dimension and on its operation. It has represented the expenditure of approximately £551,000,[790] and at least in the case of *ad hoc* evaluations, all

---

[789] The amount presented here is an estimated in GBP of the cost reported in Mexican pesos by the Ministry of Health.

[790] This is an estimated cost based on the information published by the Ministry of Health, in the case of the Design Evaluation (2007) there was no information available.

contracts have been commissioned directly to the evaluative institution. Most of these evaluations have been desk-based studies that have generated a series of findings summarised below:

- The programme is considered to have a good performance. However, it is a recurrent finding of the performance evaluation that there need to be developed more accurate indicators of the accomplishment of its objectives.
- In terms of implementation, there are some areas of improvement, especially in terms of training.
- There seems to be little evidence about the effect of the Programme in the population.
- In broader terms, it is not clear where the Programme stands within the national health quality policy.
- In terms of the *ad hoc* evaluations, the studies performed seem disconnected and do not appear to belong to a rational evaluation policy. Instead, it seems that there have been identified some key areas that have guided the performance of studies. Some of the studies seem repetitive and self-referential.

According to a Ministry of Health official (MX-3) considering the dynamics of the public sector and the model promoted by the Federal Government, 'It is difficult to perform evaluations in the way you'd like.' This reflects that the current model of evaluation is not perceived as entirely appropriate. Other interviewees supported this idea by stating that the model 'does not enable us to perform other kinds of evaluations' (MX-5) and that standardised evaluations 'are not a substitute for other kind of studies' (MX-10). These reactions about the current evaluation model are addressed more carefully in Chapter 6. However, these statements are useful for highlighting the lack of flexibility of the evaluation policy.

SICALIDAD was also audited by the SAO in 2011 as part of the revision made to the Federal Public Account. This audit comprised 'the evaluation of the efficacy in the accomplishment of objectives and goals, the efficiency in the operation of the programme, the economy in the use of resources, the progress toward the Performance Evaluation System and the accountability.'[791] The most relevant conclusions of this study can be summarised as follows:

- The SAO considers that the programme lacks effective mechanisms and indicators to verify the accomplishment of its objective.

---

[791] Supreme Audit Office. 2012. "Health Quality System Performance Audit." Mexico City.

- The SAO urges the Ministry of Health to establish mechanisms for the promotion of quality.
- The programme made rational use of the resources given for its operation.
- Most recommendations made by the SAO focused on the development of management instruments for improving operation. None of the recommendations were directed to core issues of the programme.

The findings derived from the audit to SICALIDAD are consistent with the previous discussion about the limited scope of evaluation. Both internal and external mechanisms appear to make a superficial appraisal of the programme without addressing core issues of the health-quality policy that could affect political stakes. Most of the recommendations made by the SAO to SICALIDAD were operational issues that can be easily fixed, i.e. 'to ensure the development of a supervision programme for the accredited institutions.'[792] Considering the nature of recommendations it is understandable that the SAO reports 99% of compliance on its recommendations. However, the impact of both evaluation and auditing findings on decision-making is quite limited, as a former deputy recognises (MX-23) that these represent 'no input for the Legislative.'

This overview of the case of SICALIDAD will be complemented with the description presented in the next section about the operation and evaluation of the Health Caravans Programme. This description will allow identifying those core elements of these programmes that have been central in the evaluation process. This resonates with the theoretical discussions about the control that is exercised through evaluation by limiting its scope and influencing the process. The description of the second case study from the Mexican context will expose in more detail how the Federal Government exercises this influence and the role of the external audit body in this scenario.

---

[792] Ibid.

5.3 Case D: The Health Caravans Programme

Within the Mexican health policy, the Health Caravans Programme is considered a particularly relevant policy instrument due to the objectives that are pursued through its implementation. The role of a programme the main purpose of which is providing health services to the population living in remote locations gains relevance in Mexico, where an important number of people lack effective access to health and social-care services.[793] The description of this case will show some similar features to those observed in SICALIDAD. However, the nature of the programme has given it a more visible spot in the public agenda that influences the way in which evaluation operates in a more politicised context.

5.3.1 General Overview of the Health Caravans Programme

The main objective of the Health Caravans Programme is 'to contribute to the improvement of the health conditions of the population without access to health services through offering first level attention services in medical mobile units.'[794] The programme's operation is based on a series of strategies to guide the activities towards the accomplishment of its objective (Figure 5.5). Some of its most relevant strategies are:

1. To increase the health services coverage by planning each one of the routes of the caravans; supported by itinerant health squads and by mobile medical units.
2. To establish the legal framework that regulates the operation of the programme in the states.
3. To strengthen the response capacity of the Medical Mobile Units by using technology and innovative medical and communications equipment.
4. To contribute to reducing maternal and child mortality through coordinated actions with medical attention and public health instances.
5. To develop and to link strategies for medical and dental attention, prevention, promotion, attention, and disease control of the beneficiaries of the programme.[795]

---

[793] According to CONEVAL, in 2012 the percentage of Mexican population that lack access to health services was 21.5. National Council for the Evaluation of the Social Development Policy. 2014a. "Coneval analyses the access and effective use of health services in population affiliated to *Seguro Popular* at the national level 2012." Mexico City.

[794] "Agreement that issues the operational rules of the Health Caravans Programme for the 2011 fiscal year."

[795] Ministry of Health. 2011a. "Health Caravans Programme. Specific Action Plan 2007-2012." Mexico, City.

Figure 5.5 Overview of the Operation of the Health Caravans Programme

Equipping of medical mobile units and staff's training → Integration of the existent medical mobile units → Hiring of human resources → Selection and evaluation of managerial staff → Accreditation of mobile units → Actions in prevention and promotion of health → Medical attention services
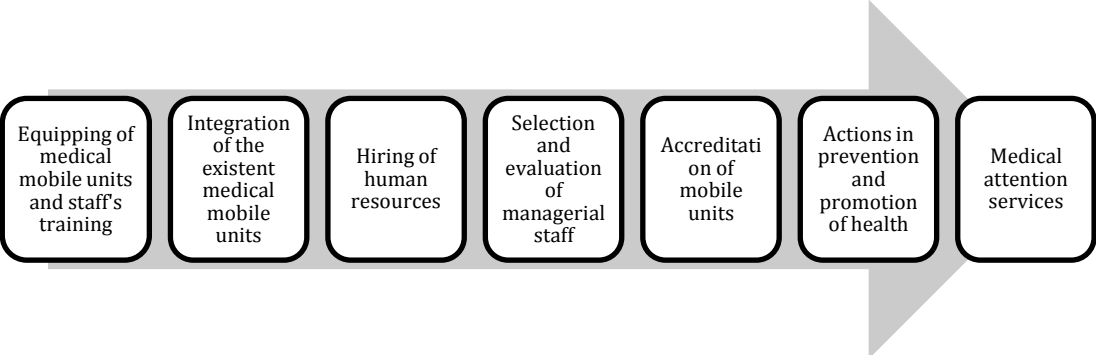
Source: Based on 2011 ORP Health Caravans.

The services provided by the Medical Mobile Units (MMU) are established in a catalogue based on the classification of the group of population that they serve, e.g. infants and children, and the kind of attention they provide, e.g. emergencies. This catalogue establishes a more structured organisation for the activities developed by the programme. In addition, the programme has established a set of indicators to monitor its main activities, also reported to coordination agencies (Table 5.6).

Table 5.6 Indicators of the Health Caravans Programme

Name	Periodicity
Percentage of people served by the programme in relation to the total population living in isolated communities with high levels of marginalisation	Annual
Percentage of communities served in relation to disperse communities that are difficult to access	Annual
Percentage of promotion and prevention services carried out in relation to those planned	Biannual
Percentage of ambulatory medical services provided in relation to those planned	Biannual
Percentage of medical mobile units fully equipped and in operation that have complete and trained itinerant team	Quarterly
Percentage of health staff specifically trained for the programme	Quarterly

Source: Based on ORP Health Caravans 2012.

As in the case of SICALIDAD, the use of indicators for monitoring the activities oversimplifies the potential effect of the programme. The complexity of the problem addressed by Health Caravans (the lack of access to health services) requires the use of more sophisticated approaches to evaluate its effects on the target population and to determine which changes need to be made for its improvement. This complexity is also perceivable in terms of the coordination

level that the programme demands. As an official of the Ministry of Health (MX-2) states: 'If the programme requires the intervention of other areas to achieve its full potential, you don't get the results you want because of the coordination and participation difficulties.' The different conceptualisations about how a programme is evaluated affect the necessary conditions for fostering the use of evaluation findings. The following section more deeply examines the evaluation process of the Health Caravans Programme which provides the context for the analysis of the hidden politics of evaluation in Chapter 6. As it is possible to observe in this case that there is a clear bias of evaluation towards efficiency and effectiveness, indicators are underpinned by this rationale and the instruments promoted by the Federal Government are directed to generate information about these values.

5.3.2 The Evaluation of the Health Caravans Programme

The Health Caravans Programme has also been evaluated systematically since 2007 with different instruments that provide information about its performance (see Table 5.7 below). All evaluations have been performed under the scheme promoted by the Federal Government and the SAO has only audited the programme in the context of the revision of the Federal Public Account in certain states of the country. It is important to acknowledge the kind of evaluations performed to identify which areas of analysis have been prioritised. Most mandatory evaluations have focused on the performance of the programme using the standardised frameworks developed by CONEVAL, while *ad hoc* evaluations have analysed more specific issues derived from the mandatory studies, such as the coverage of the programme.

Table 5.7 Evaluations Performed on the Health Caravans Programme 2007-2012

Year	Evaluation Type	Evaluator	Budget[796]	General Description of the Study
2007	Design Evaluation	Demography and Health Research, S.C.	Not Available	It assesses the consistency on the design of the Programme through a standardized framework developed by CONEVAL.
2007	Complementary Evaluation	Metropolitan Autonomous University	£167,000	It was a diagnosis of the access to health services for the target population of the Programme.
2008	Complementary Evaluation	Beta-Korosi, S.C. (Private consultant)	£90,000	It was a study about the coverage and operation of the Programme, as well as its articulation with other health programmes, involving both documentary analysis and fieldwork.
2008	Specific Performance Evaluation	Demography and Health Research, S.C.	£3,800	It is an annual synthetic review of the performance of the programme in terms of a series of management and result indicators, as well as its budget performance. It aims to identify key information about the performance of the Programme through the analysis of its indicators. It is based on a standardized framework developed by CONEVAL.
2009 - 2010	Specific Performance Evaluation	National Institute of Public Health	£3,100	
2010 - 2011	Specific Performance Evaluation	Dr. Antonio Carlos Martín del Campo Rodríguez (Independent consultant)	£3,000	
2009	Complementary Evaluation	Meritorious Autonomous University of Puebla	£171,000	It was a study focused on the results obtained by the Programme in issues like coverage and operation, involving both documentary analysis and fieldwork.
2010	Complementary Evaluation	National Autonomous University of Mexico	£139,000	It was a cross-sectional study that compared people in the programme to other people in similar conditions that don't have access to it.
2011	Complementary Evaluation	National Autonomous University of Mexico	£255,000	It assessed the coverage of the Programme in locations with very low HDI, it also measured user's satisfaction and improvement in the conditions of the target population.
2012	Complementary Evaluation	National Autonomous University of Mexico	Not available	It was a meta-evaluation that comprises the reports delivered during 2007-2011.
2011 - 2012	Consistency and Results Evaluation	Dr. Nelly Aguilera (Independent Consultant)	£15,000	It is a desk-based study that aims to analyse if programmes have the necessary information for its operation, how it produces new information, and how this information is used for the improvement. Based on a standardized framework developed by CONEVAL.

[796] The amount presented here is an estimated in GBP of the cost reported in Mexican pesos by the Ministry of Health.

Table 5.7   *Continued*

Year	Evaluation Type	Evaluator	Budget	General Description of the Study
2012-2013	Specific Performance Evaluation	National Institute of Public Health	£4,000	The Programme has reached its planned coverage; however, its coverage is based on weak definitions and quantifications of the population. The lack of a list of beneficiaries makes difficult the identification of the distribution of the attention given to the population. It is necessary to generate more evidence that allows identifying the real access to the services offered.
2013	Complementary Evaluation	Intercultural University of the State of Mexico	£421,000	The study assesses the effects of the Programme in indigenous locations with high and very high levels of marginalization.

Source: Based on data from the Ministry of Health.

These evaluations (mostly desk-based studies) represent the expenditure of approximately £1,271,900.[797] These have produced evidence about some areas of improvement for the programme, for instance:

- The coverage of the programme needs to be increased
- Its performance is adequate. However, there are some operational issues that can be improved
- There is little evidence about the impact of the programme on the target population.

The evaluation of Health Caravans follows the evaluation agenda established centrally by the Federal Government. Also the process for the follow-up of recommendations is promoted and conducted by this actor. Consistently with the case of SICALIDAD, there is a very narrow space for programme operators to identify, design, and implement evaluation instruments beyond those already established. As a governmental initiative, Health Caravans represents an example of a programme based on a highly complex operation that relies on the coordination of multiple participants; a complexity that is also transferred to the evaluation arena. Taking into consideration the nature and purpose of the programme, it appears that the evaluation process undertaken has generated little information about the overall effect of the programme in terms of the access to health services. Instead, the focus has been on operational elements of the programme, which, according to an operator (MX-11) 'have not been useful for generating new information.' From this perspective, the questionings about the model of evaluation show that the interests of stakeholders within the Federal Government are not necessarily aligned, and consequently this can generate disputes among them.

This chapter has described the institutional framework for evaluation and audit at the national level in Mexico. From this description it is possible to make some relevant remarks for the analysis of the Mexican case. Regarding internal evaluation, the intricate legal framework generates complexity and ambiguity about the competences of coordination agencies. At the same time, the set of

---

[797] This is an estimate based on information available on the Ministry of Health and CONEVAL's websites and it includes the cost of the evaluations performed to the Health Caravans Programme for the 2007-2012 period with exemption of the Design Evaluation 2007 and the Complementary Evaluation performed in 2012 where there was no information available.

rules and institutions involved in evaluation give the Federal Government an important amount of power that places other stakeholders, e.g. operators, in a disadvantaged position for participating in the evaluation process.

In relation to auditing, the role of the SAO as the organisation in charge of making the Federal Government accountable is weak due to its lack of institutional capacities to promote changes in public organisations and to sanction the inobservance of the law. Moreover, these two processes coexist but do not integrate into a more comprehensive evaluation system that could foster the impact of these activities in policymaking.

This characterisation of the Mexican case becomes relevant in the light of the analysis of the hidden politics of evaluation for two main reasons. On the one hand, the uneven allocation of power will affect the dynamics that are created around evaluation. On the other hand, the influence that a particular stakeholder (in this case the Federal Government) can exercise over the process will impact the outputs and outcomes of evaluation according to the interests, values, and beliefs held by this actor.

The most relevant feature is the complexity associated with the multiple actors involved, the control exercised by the Executive, and the lack of political power of the audit body. Consequently, in the Mexican case a fragmented evaluation policy is observable, a weak audit procedure and a lack of enforcement capacity of the recommendations derived from both instruments. In the specific cases of the SICALIDAD and Health Caravans Programmes, the rigidity of the evaluation instruments applied allows very little space for a more sophisticated and plural discussion about their effectiveness in relation to the problem that originated their design and implementation.

The next chapter dissects the evaluation and audit processes of all four case studies described through the lenses of the analytical framework for the understanding of the hidden politics of evaluation (see Chapter 2).

Chapter 6. A Comparative Analysis

Previous chapters have described the institutional framework in which evaluation and audit processes take place in the UK and Mexico (see Chapters 4 and 5). The description about the rules and institutions that shape evaluation and audit in both countries and the results of semi-structured interviews with relevant stakeholders have generated empirical evidence about the hidden politics of evaluation. The collection and systematisation of this data would be a sterile effort if the researcher were not to go through an analytical process to support the hypothesis of this research. This chapter now explores the interaction between evaluation and politics and how it is affected by the institutional framework. Therefore, its purpose is to identify, compare, and contrast the most relevant findings of the documentary and empirical research through the use of the thematic framework proposed in Chapter 2. The institutional framework is the core variable of the analysis, both for the general findings as well as the specificities of each case study (Figure 6.1).

Figure 6.1 General Outline of the Analysis
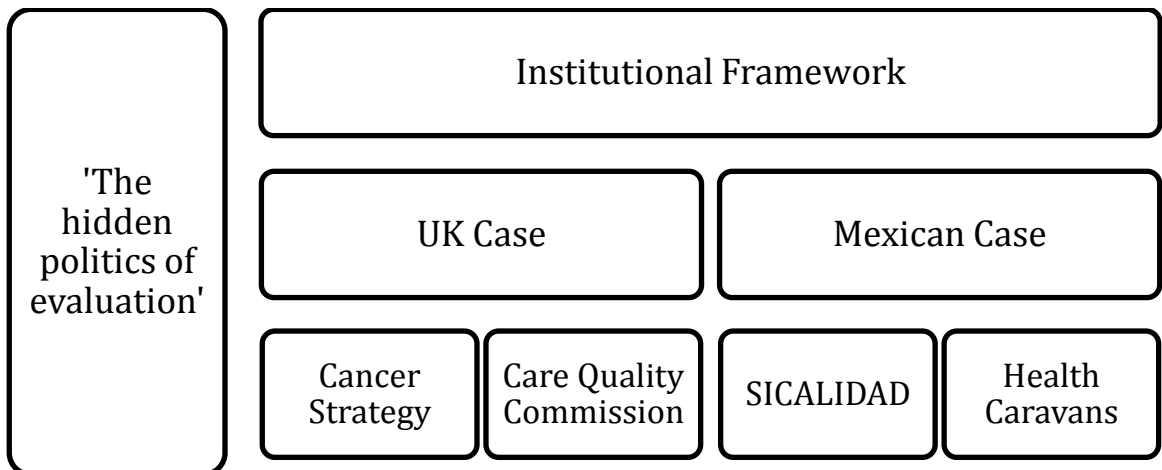
The core argument of this chapter is that the institutional framework, i.e. rules, coordination mechanisms, power distribution, etc., *does play* an important role in the politicisation of evaluation. The multiple configurations of power and resources built around evaluation based on formal and informal arrangements impact how this activity interacts with politics. The nuances of this argument concern the level of control that can be exercised over the evaluation space and

the resources that stakeholders have at their disposal for this purpose.

The chapter is divided into three main sections. The first analyses the findings of the UK case through the evidence derived from the case studies (Cancer Strategy and the CQC). The second section replicates this analysis for the Mexican case studies (Health-Quality System Programme and Health Caravans Programme). The last section compares and contrasts both countries to make explicit the linkage between the empirical evidence and the theoretical framework.

6.1 Main Findings of the UK Case

Different monitoring and evaluation mechanisms have been implemented in the UK. In the public sector, there are internalised practices that systematically produce information about policy through the application of various methods (see Chapter 4, Section 4.1). In terms of monitoring, for example, the UK possesses a system controlled by HM Treasury that collects data from government departments about the accomplishment of organisational goals (indicators). Policy evaluation is underpinned by a different logic. There is neither a formal central mechanism to enforce policy evaluation nor a standardised process for its performance. Instead departments have guidelines at their disposal for evaluating their policies. However, in practice every organisation has its own routines and procedures.

The main strength identified in the UK's institutional framework is not related to evaluation but to auditing, particularly to the work of the NAO. The accountability system relies on the role of this body and the follow-up inquiries carried out by Select Committees of the House of Commons. Neither the NAO nor the Select Committees have faculties to question the merits of public policy. They are statutorily prohibited from doing so; *ergo,* their work is basically limited to the analysis of operational issues. This is an important element to highlight as the NAO's role (and therefore PAC's) is producing and debating evidence about policy effectiveness, not questioning the government's solution to a particular problem. The question that arises here is: *To what extent is it is possible to respect the thin line that stands between what is merely procedural*

236

and what is not?

These three main elements of the UK's institutional framework can be summarised as follows:

- A performance-monitoring system operated by HM Treasury that generates inputs for the budgetary system. However, its actual utilisation has been largely questioned (see Chapter 2, Section 2.1.3)
- A heterogeneous evaluation policy (subject to each department's procedures) whose purpose is generating knowledge for the internal decision-making process
- A strong audit system guided by the values of efficiency, efficacy, and economy

This brief description of the functioning of the institutional framework for policy evaluation in the UK exposes the conditions that shape the behaviour of stakeholders and their interaction, e.g. between commissioners and evaluators.

6.1.1 The Institutional Framework and the Hidden Politics of Evaluation in the UK

The hidden politics of evaluation concerns the establishment of relationships between stakeholders and how they use their resources, i.e. information, power, money, etc., to defend their interests. The analysis of this phenomenon in the UK case is guided by the key political elements involved in evaluation and the characteristics identified in the institutional framework (Figure 6.2).

Figure 6.2 Main Characteristics of the Institutional Framework for Evaluation in the UK
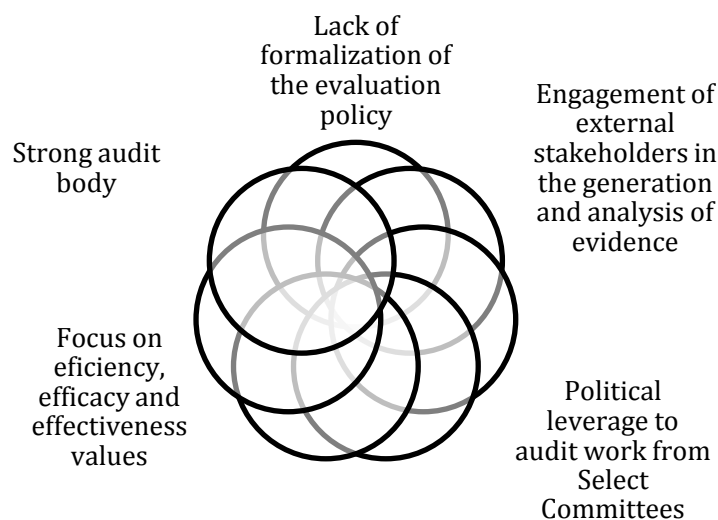
The lack of formal mechanisms of enforcement for evaluation in the UK has turned this activity into an instrument for internal decision-making and organisational learning. The importance of a formal evaluation policy has been explained by Trochim, who underlines the need for explicit criteria to guide the development of this activity.[798] This absence of formality has developed into an uneven construction of evaluation capabilities within central government. The DfID, for example, stands out because it has devoted important efforts to the establishment of an evaluation policy 'to determine whether the desired effects are being realised.'[799] This case is an outlier in the evaluation landscape, as no other department has developed an evaluation policy to this degree, not only in terms of the use given to evaluation findings but also because of the recognition of the 'progress made and the challenges and opportunities encountered in embedding evaluation across the organisation.'[800]

These institutional conditions have given policymakers plenty of freedom to use evaluation more as an aid for internal decision-making than for the legitimisation of public policy. This gives evaluation a modest role in the broader context of governance and an unclear effect on policymaking. The Institute for Government has recognised this by claiming that the government needs to play …

> … a greater role for the centre in overseeing the quality of policy making (rather than just skills and capabilities) through the creation of a senior Head of Policy Effectiveness, who will also ensure rigorous and independent evaluation of government policies, and commission lessons learned exercises for major failures of policy process.[801]

Nonetheless, this lack of standardisation has an important advantage. Under these conditions, central departments are able to evaluate considering their particular needs, using the methods identified as most suitable and with the periodicity believed as appropriate. However, this produces fragmentation for

---

[798] Trochim, W. M. K. 2009. "Evaluation policy and evaluation practice." *New Directions for Evaluation* 2009(123):13-32.

[799] Department for International Development. 2013. "International Development Evaluation Policy." London, UK. P. 3.

[800] UK Aid, and Department for International Development. 2014. "Rapid Review of Embedding Evaluation in UK Department for International Development." London, UK. P. i.

[801] Rutter, J., and M. Hallsworth. 2011. "Making Policy Better. Improving Whitehall's core business." London, UK: Institute for Government. P. 9.

the consolidation of an integral evaluation policy. Outputs are dispersed and do not provide a straightforward judgement about the effectiveness of public policy.[802]

The UK case shows that there is an important production and use of evidence by different organisations within and outside the government sphere. To illustrate, between 2011 and 2014, the DoH published circa 60 research and analytical papers with robust evidence about a wide range of topics related to health care.[803] Multiple actors were involved in the generation of evidence, i.e. PHE, NICE, and NGOs, and they also participated in the inquiries performed by Select Committees.[804]

The level of influence of each stakeholder is an important variable to consider in terms of the hidden politics of evaluation. In the case of NGOs, for example, it is possible to observe asymmetries in the level of engagement among these organisations. This might be related to the resources at their disposal and their lobbying capacity.[805] Cancer Research UK, for instance, is one of the most active NGOs in health policy. This organisation has provided evidence in numerous Select Committee inquiries [806] and it has also promoted Early Day Motions[807] in Parliament for the discussion of diverse topics related to cancer, e.g. obtaining funding from the government.[808] This evinces there being channels for stakeholders to intervene in the discussion of public issues and to have a more

---

[802] See National Audit Office. 2013a. "Evaluation in government." edited by National Audit Office. London, UK.

[803] Some recent examples are the *Cancer Survivors Survey: Quality of Life after Diagnosis and Treatment* and the *Study on the Impact of Digital Technology in Health and Social Care*, both published in 2013.

[804] As the former National Clinical Director for Cancer and End of Life Care recognised 'the guidance that NICE has produced, the peer review programme, all of those collectively have helped us to make progress on cancer.' Committee of Public Accounts. 2011. "Delivering the Cancer Reform Strategy." London, UK. PAC Report. Ev. 1.

[805] There are multiple NGOs that are importantly engaged in the discussion of public affairs, for instance, Oxfam and Action Aid.

[806] For example: The Clinical Trials Inquiry made by the Science and Technology Committee (2013) of the House of Commons and the Tobacco Products Directive inquiry led by the Select Committee on the European Union of the House of Lords (2013).

[807] These are 'are formal motions submitted for debate in the House of Commons.' Parliament UK. "What are Early day motions? ." London, UK.

[808] See: —. 2013. "Early Day Motion 1124 Cancer Research UK." London, UK.

active participation in this process.[809]

Despite this evidence-based approach, a gap is perceivable between the utilisation of the evidence produced and the context in which policy takes place.[810] As the theory states, a key component of evaluation is issuing a judgement about an object, considering previously set criteria. In the absence of this element (judgement), evidence becomes part of the information at the disposal of stakeholders to strengthen the decision-making process, but it cannot be considered as 'evaluation' because it has not gone through an analytical process by which outcome can be used to determine the value of an object.

The lack of a structured evaluation policy has generated, as the NAO recognises (see Chapter 4, Section 4.1.2) heterogeneity in the quality of evaluation and little knowledge about the general state of this activity in central government. In terms of the institutional framework and its relationship with politics, elements like centralisation, distribution of power, and levels of discretion are crucial in the discussion. In terms of centralisation, the role of HM Treasury is modest. It has developed the guidelines for this activity (*The Green Book* and *The Magenta Book),* but it does not oversee the application of these instruments. Although these guidelines do not establish formal obligations, they do make evident the logic behind the use of evaluation in government:

> The Treasury has, for many years, provided guidance to other public sector bodies on how proposals should be appraised, before significant funds are committed – and how past and present activities should be evaluated.[811]

This reveals the connection between the use of evaluation tools and the improvement of decision-making for the allocation of resources. On the one

---

809 For example, in the recent study about dementia was conducted by the Senior Regulatory Officer, Integrated Development, Global Health, Bill & Melinda Gates Foundation and it also incorporated the views of different stakeholders, e.g. OECD and World Health Organization, for both the conduction and reviewing of the report. Long, R. 2015. "Finding a Path for the Cure for Dementia. An independent report into an integrated approach to dementia research.". London, UK: Department of Health.
810 See: Nutley, S., H. Davies, and I. Walter. 2002. "Evidence Based Policy and Practice: Cross Sector Lessons From the UK." in *Social Policy Research and Evaluation Conference*. Wellington, New Zealand.
811 HM Treasury. 2011a. "The Green Book. Appraisal and Evaluation in Central Government." London, UK. P. v.

hand, there is an important concern in having more tools in reserve to make a more efficient use of public funds. On the other hand, it reflects the need for evidence about policy effectiveness. The allocation of resources is a political decision that requires technical support to convince others of the pertinence of spending more or less money in a policy area, particularly when there are disputes about who gets more funding and who will be affected by budget cuts.[812] The second element is also political, as it regards the identification of the government's effectiveness and, in consequence, the legitimisation of public action (see Chapter 2, Section 2.3.1). For example, the relevance given to impact evaluation can be explained because this tool 'supports both accountability, by verifying whether what was promised was also delivered, and learning, by drawing from the experience in order to better design and manage future interventions.'[813] The recognition of the effectiveness of a particular policy can be translated into political backing for a politician or a political party. It shows the symbiotic relationship between the rational and political dimensions of evaluation as 'technical expertise and facilitation of knowledge access would be complemented by an impact feedback on political decisions.'[814]

The most distinctive feature of the UK case, therefore, is the absence of formal rules and obligations in evaluation matters. This can be interpreted as a mechanism of evasion in terms of accountability. Without formal rules and expectations for evaluation, it is difficult to determine to what extent this activity is embedded in the public sector and, more importantly, what has been the effect of performing evaluations. Also it impedes determining with clarity how stakeholders will interact and what is expected from these relationships. Adopting a more formal model of evaluation would imply establishing responsibilities for the coordination of this activity, e.g. to HM Treasury. This was acknowledged by the Institute of Government:

---

[812] For the Comprehensive Spending Review 2015-2016 presented by the Chancellor of the Exchequer, George Osborne MP, there were important cuts in different policy areas and substantial increases in some others like the NHS. See: 2013b. "Osborne's spending review – winners and losers at a glance " *The Guardian.*

[813] Scott, A. Ibid."Impact evaluation: which way forward? ."

[814] Steinich, M. 2000. "Monitoring and Evaluating Support to Decentralisation: Challenges and Dilemmas." European Centrre for Development Policy Management. P. 7.

> The Treasury ought to play a key role in incentivising departments to commission and use good evaluations – but that it probably played that role less than it should, perhaps because Treasury spending teams spent more time mediating between competing departmental priorities than focusing on the cost effectiveness of what money was spent on.[815]

This is also linked to its role in the development and refinement of evaluation guidelines. Neither *The Magenta* nor *The Green Book* have been recently updated. Current editions were published in 2003, with some additions in 2013 and in 2011, respectively. This leads to reflecting upon: *To what extent was policy evaluation a core element of the Coalition's government agenda?* At least in terms of the establishment or follow-up of formal guidelines, there were not substantial efforts during this period. Without a structured guidance for evaluation it is not possible to determine whether departments should be producing more or less evaluations; if more sophisticated studies should be prioritised in contrast to more policy-oriented outputs; or even if the model of commissioning evaluations is more pertinent than an in-house approach. This ambiguity leaves space for different interpretations about the role of evaluation in policymaking and also about the values that policymakers associate with evaluation as these 'influence not only how they make their decisions but also the content of those decisions.'[816] As a stage of the policy process, evaluation should represent a mechanism to identify the effects of an intervention and to contrast them with the expected results identified in the design. But as an instrument for accountability, evaluation demands a less instrumental perspective to discuss the effectiveness of a government and its capacity to respond to social demands in a context in which the meaning of effectiveness is built on the perceptions of different stakeholders. Consequently, the role of evaluation in the UK is unclear. Efforts in this matter are dispersed and fragmented and understanding its purpose as a governance tool is not a straightforward process.

In terms of auditing, it is possible to identify two major purposes. The first one is to make central government accountable in relation to the VFM that

---

[815] Rutter, J. 2012. "Evidence and Evaluation in Policy Making. A problem of supply or demand?", edited by Institute for Government. P. 21.

[816] Peters, B. G. 2001. *The Politics of Bureaucracy*. London, UK: Routledge. P. 43.

departments deliver through their policies. The second concerns the power to insert in the public agenda different policy issues relevant to the electorate or for groups of interest. This role of leverage played by audit bodies provides a more democratic context for questioning the effectiveness of government, particularly when civil society is involved.[817] The institutional framework sets the legal obligation for the NAO to make central government accountable, both by scrutinising public accounts as well as through the performance of VFM studies. This comprises the support of the PAC to the NAO's work, giving auditing a very important political leverage.

The UK case offers an interesting example of a context in which the institutional framework serves as a mechanism to limit the power of the NAO in the aim of preserving independence and objectivity. The work of this body is underpinned by three main values:

> *Accountability* – everything done by those who work in the NAO must be able to stand the test of parliamentary scrutiny, public judgements on propriety, and professional codes of conduct
> *Probity* – there should be an absolute standard of honesty and integrity in handling NAO work and resources, and
> *Objectivity and Impartiality* – The C&AG's independence is secured in statute. This underlines the need for us to be objective and impartial in all our work, including accurate, fair, and balanced reporting.[818]

This is relevant at a discursive level. However, the way in which the legal framework sets a basis for the NAO's operation provides safeguards for the independence and objectivity of its judgements. This is how the institutional framework can impact the way in which politics operate, when the rules and the distribution of power can shape the behaviour of a stakeholder. In the case of the NAO, its institutional architecture imposes a safe distance between the NAO and the Executive (through PAC's intermediation). As a member of this Committee (UK-12) recognises:

> In our culture and in the standing orders of the House of Commons, the Committee of Public Accounts […] must be chaired by a member of the opposition, […] the Comptroller Auditor General when appointed is an

[817] Malena, C., and M. McNeil. 2010. "Social Accountability in Africa: An Introduction." in *Demanding Good Governance: Lessons from Social Accountability Initiatives in Africa*, edited by C. Malena and M. McNeil. Washington, D.C., US: The World Bank.
[818] National Audit Office. "Quality and Independence." London, UK.

> independent officer of the House of Commons and then the budget [...] comes from Parliament, that's again a very important safeguard, so we have a separate body called the Public Accounts Commission, [...] the job of which is to listen to the NAO's bid for a budget and then approve it, and then sign it off, we give the information to Treasury, they write the cheque but they write the cheque of the size that we say, they have no say over the budget of the NAO, again that provides an extra distance and extra institutional independence ...

These institutional safeguards endow a more independent space for the NAO to scrutinise governmental bodies, while at the same time constrain it from exceeding its functions through the surveillance of the PAC. Although this Committee does not follow up all NAO reports to the same extent, it does systematically monitor its activity, giving more legitimacy to its work. This leads to another important element of analysis in relation to politics. The reports delivered by the NAO and followed up by the PAC might reflect those policy areas considered more relevant for MPs. For instance, during 2014 the NAO published 70 major reports, and nearly 47% were followed up by the PAC.[819] Derived from this, it comes to our attention that health, infrastructure, and education are high-priority issues on the Committee's agenda.[820] It is a political decision to select and prioritise the evaluation or auditing of a particular policy or programme. In the absence of a policy that dictates key elements of the process of evaluation, i.e. periodicity, scope, this decision becomes discretional. This is consistent with what Weaver and Brockman discuss about the *trade-offs* of different institutional arrangements.[821] While the NAO sits distantly from the government, it is immersed at the same time in the political logic of the PAC.

The UK case is also characterised by the existence of mechanisms for the participation of both governmental and non-governmental stakeholders in the audit process, i.e. VFM studies, or Select Committee inquires. In a sensitive area, such as public health, groups of interest and civil society need formal mechanisms to participate in policymaking, and yet this does not exempt it from

---

[819] This is an estimate based on the number of reports published by the NAO during 2014 and the number of follow-up reports delivered by the PAC during the current Parliament. Sources: Committee of Public Accounts. 2015b. "List of Reports from the Committee during the current Parliament." London, UK: House of Commons, National Audit Office. "Major Outputs." London, UK.

[820] See: Committee of Public Accounts. 2014c. "Publications." London, UK.

[821] Weaver, K. R., and B. A. Rockman. 1993. *Do Institutions Matter? Government Capabilities in the United States and Abroad*. Washington, D.C., US: The Brookings Institution.

politics because it might involve a close—sometimes adversarial—relationship between the Executive and Parliament. As Gray and Jenkins observe:

> When faced by critical reports from bodies such as the NAO and the Audit Commission, the executive in the United Kingdom has attempted to narrow the parameters of what it is and should be accountable for [...] at the centre accountability is often defined narrowly, i.e. in a managerial or technical fashion, with the function of audit restricted to serve this.[822]

The variety of stakeholders involved in these processes generates more inclusive conditions as the incorporation of other views contributes—at least formally—to reducing the risk of capture and bias in the inquiry. To illustrate, the PAC's inquiry about adult social care in England brought in the oral-session witnesses from different sectors whose perspectives were later included in the report. NGOs such as the Association of Directors of Adult Social Services, Carers UK, and Age UK participated along with DoH and the Department for Communities and Local Government Staff.[823] This scheme provides at least the opportunity for external stakeholders to raise their concerns about a particular topic, and to have their views represented by the actor leading the process.[824] In this case, for example, the Association of Directors of Adult Social Services stated that 'Collaboration is needed across services beyond health and social care, within both central and local government', while Carers UK added that 'Changes to housing benefit may mean that when people with disabilities have to move house, their caring relationships could suffer as a result.'[825] Although it is not possible to determine the extent to which the participation of these actors effectively influences policymaking, the institutional framework allows two major issues:

- It permits external stakeholders to share their views and to make the PAC aware of their concerns.

---

[822] Gray, A. , and B. Jenkins. 1993. "Separate Developments: Budgeting, Auditing, and Evaluation in the United Kingdom." in *Budgeting, Auditing, and Evaluation. Functions and Integration in Seven Governments* edited by A. Gray, B. Jenkins, and B. Segsworth. London, UK: Transaction Publishers. P. 59.

[823] Committee of Public Accounts. 2014a. "Adult social care in England." London, UK: House of Commons.

[824] Mathison, S. 2000. "Deliberation, Evaluation, and Democracy." *New Directions for Evaluation* (85):85-89.

[825] Committee of Public Accounts. 2014a. "Adult social care in England." London, UK: House of Commons. P. 11.

- The incorporation of these views in the reports provides more leverage for stakeholders to pursue their interests.

Select Committee inquiries are politically relevant because of the effect that these might have in the decision-making. For example, Benton and Russell recall the case of the Health Committee's *Smoking in Public Places* inquiry that 'provided crucial political reinforcement for the mounting medical evidence, and clearly influenced other MPs' views.'[826] Nonetheless, the participation of stakeholders in this process is mostly limited to the provision of evidence which is based on formal guidelines established for this purpose.[827] There are open calls for evidence, when any person interested can easily identify the inquiries taking place and submit evidence.[828] However, this formality does not completely eliminate the spaces of discretion. As a Health Committee staff member (UK-2) recognises, this procedure is constrained by the resources available in terms of staff and time:

> I'd read the evidence and I take what I think are the important points and I put that into our briefings and there's no, there's no formal fact [...] I wouldn't have the resource or the time to go away and double-check all these claims. In no way I can do it; you have to take them at face value.

This suggests that actors who control the inquiry process have the prerogative to include or dismiss a particular piece of evidence based on their own criteria. This is also affected by the barriers of access. Stakeholders with a more public profile and more resources would have more elements to participate and defend their interests.[829]

---

[826] Benton, M., and M. Russell. 2013. "Assessing the Impact of Parliamentary Oversight Committees: The Select Committees in the British House of Commons." *Parliamentary Affairs* 66(4):772-97. P. 787.

[827] In the case of civil servants, for instance, 'ministers have emphasised that, when officials represent them before Select Committees, they should be as forthcoming and helpful as they can in providing information relevant to Committee inquiries.' Cabinet Office. 2014b. "Giving Evidence to Select Committees. Guidance for Civil Servants." P. 3. For members of the public, there is also guidance for this process. See: Parliament UK. "Guidance on giving evidence to a Select Committee of the House of Commons ". London, UK: House of Commons.

[828] Currently, there are open calls for evidence for inquiries undertaken by the Northern Ireland Affairs Committee and the Science and Technology Committee. See: —. "Committee inquiries: Open calls for evidence." London, UK: House of Commons.

[829] This can be related to the issue of lobbying as a way of searching mechanisms to influence decision (e.g. to get an MP to vote in a certain way), the extent to which individuals or organisations can influence depends on multiple factors i.e. resources). According to the UK

This general overview of the hidden politics of evaluation in the UK serves as a point of departure for understanding how different relationships of power and interchange of resources take place in this context. The following two sections focus on the case studies (Cancer Strategy and the CQC) which are dissected with the aid of the thematic framework proposed (see Chapter 2).

6.1.2 The Hidden Politics of Evaluation in the Cancer Strategy

The description of the Cancer Strategy (see Chapter 4) has shown the complexity that characterises public policy when it involves the participation of multiple stakeholders, not only in its design and implementation but also in the way in which the meaning of success is built. The Cancer Strategy is a policy instrument directed to address different purposes. On the one hand, it aims to improve the outcomes related to this disease, i.e. improve survival rates or reduce emergency admissions; on the other hand, to increase the focus of the government's intervention on the patient's needs and the quality of the services offered. The definition of whether the Cancer Strategy has been successful or not depends on the actor evaluating it and the criteria used for this purpose. In general terms, the evaluation and auditing of the Cancer Strategy do not provide a clear statement about the results obtained by this policy. However, the information produced by these processes does give important insights about the hidden politics of evaluation.

The first element to highlight is the evidence-based nature of this case. The proliferation of scientific evidence about cancer plays an important role supporting decision-making. Cancer is considered a technical subject, but simultaneously it represents a highly sensitive policy issue. Its scientific nature helps reducing the controversy around decisions taken, as the evidence on which it is based is considered legitimate. For example, the target of saving an extra 5,000 lives per year represents an important political statement, as claimed by the former Secretary of State for Health:

Public Affairs Council (UKPAC) lobbying 'means, in a professional capacity, attempting to influence, or advising those who wish to influence, the UK Government, Parliament, the devolved legislatures or administrations, regional or local government or other public bodies on any matter within their competence.' UK Public Affairs Council. 2012. "Lobbying Definition."

> The coalition government's reforms of health and care services will drive improvements in what matters most to patient and their families – cancer outcomes, lives saved. That is what we will be measuring our success against.[830]

The support given to the statement is based on the potential benefits identified by the scientific community of implementing a series of actions, i.e. early diagnosis. In addition to the political claim of considering cancer a top government priority, the scientific support given to these decisions contributes to their legitimacy.[831] Moreover, the incidence of this disease among the population inevitably gives cancer a major spot in the public agenda, which cannot be ignored by politicians.[832] To illustrate, the recent election promoted the debate about cancer and the government's response.[833] Different media, for example, reported the flaws of the Coalition Government in this area:

> More than a third of cancer patients are dying within a year of being diagnosed in parts of Britain in what campaigners have described as an 'inexcusable' postcode lottery.[834]

> 'Hidden at Home – the social care needs of people with cancer' reveals that one in 10 (11%) people with cancer in the UK, equivalent to at least 160 000 people, say they are constantly or often left housebound due to a lack of support. At least 100 000 people (7%) are constantly or often unable to wash themselves, dress or go to the toilet […] Macmillan is urgently calling on the NHS and local authorities to recognise that people with cancer do have social care needs and they desperately need more support.[835]

> Nine in 10 people think that cancer screening is 'almost always a good idea' despite the fact that screening uptake is lower […] In the lead up to this year's General Election Cancer Research UK is calling on candidates from all parties to back measures to improve early diagnosis of cancer […] the Government should ensure accurate information is provided on the risks and benefits of all cancer screening to help people make informed decisions alongside specific efforts to increase the low uptake of bowel screening.[836]

---

[830] Watt, N. 2011. "New GP cancer tests will save 5,000 lives a year, says Lansley." in *The Guardian*. London, UK.

[831] For example, the rise in the funding for cancer in the UK represents a way to make a statement about the relevance given to this issue. See: Cancer Research UK. 2015a. "£14 million Government investment to help 'personalise' treatment." London, UK.

[832] Boseley, S. "Half of people in Britain born after 1960 will get cancer, study shows " in *The Guardian*. London, UK.

[833] The way in which cancer can be treated as a health policy issue has been studied by different scholars. For example: Faguet, G. B. 2008. *The war on cancer: An anatomy of failure, a blueprint for the future*. London, UK: Springer.

[834] Wheeler, C. "EXCLUSIVE: Scandal of postcode lottery costing cancer patients' lives." in *Express*. London, UK.

[835] McMillan Cancer Support. "Lack of social care causing devastating consequences for 100000s of cancer patients."

[836] Cancer Research UK. "Ninety per cent approve of cancer screening but screening uptake is

This evinces that cancer is a highly political policy issue which also affects its evaluation and auditing.[837] The Cancer Strategy is complex in terms of the different actors involved in its operation and the narrow scope of its assessment, mostly based on the values of efficiency, efficacy, and effectiveness. To obtain a fine-grained analysis of the hidden politics of evaluation in this case it will be used as the thematic framework developed in Chapter 2 (Figure 6.3).

Figure 6.3. Overview of the Hidden Politics of Evaluation in the Cancer Strategy

Cancer Strategy	**Purpose**
	Focus on VFM and the identification of quantifiable variables
	Predetermined and controlled by audit bodies (NAO, PAC)
	Exclusion of important values from the assessment (i.e. access, equality)
	Resources
	Important amount of data and specialists in the field
	Control of data by key stakeholders
	Legitimacy provided by academics
	Process
	Involvement of the evaluated body in the design of evaluation instruments
	Asymmetry in the level of influence of stakeholders
	Outputs
	Discretion in the follow-up recommendations
	Tensions about who gets recognition for changes in the strategy
	Discretionary dissemination of findings
	Outcomes
	Difficult to isolate the effect of evaluation in changes in the strategy
	Evidence-based but highly political policy issue

In line with the literature, the *purpose* of the evaluation establishes the premises for the whole process and it can predetermine the results.[838] In the Cancer Strategy, the internal evaluation process (performance report issued by the DoH, NHS, and PHE) clearly establishes the purpose and the criteria used for evaluating progress. The claim about saving 5,000 lives per year, for example, is a highlighted aspect in the report:

> For the first time, we have estimates of how many additional lives may have been saved against our ambition of saving an additional 5,000 lives a year by 2014/15, to halve the gap between the survival estimates in England and those in the best countries in Europe.[839]

---

lower." London, UK.

[837] See for example: Estrich, S. 2013. "The Politics of Cancer.", Pym, H. 2015. "Cancer, politics and public money - a difficult mix."

[838] Bogenschneider, K., and T. J. Corbett. 2010. *Evidence-Based Policymaking: Insights from Policy-Minded Researchers and Research-Minded Policymakers.* New York, US: Routledge, Palumbo, D. J. 1987. *The Politics of Program Evaluation.* California, US: SAGE.

[839] Department of Health, Public Health England, and NHS England. 2014. "Improving Outcomes:

As the stakeholders in charge of verifying progress, the DoH, the NHS, and PHE have the prerogative to establish and set the scope of the analysis by focusing on those areas of the government's response to cancer that they consider relevant or politically profitable, especially those quantifiable:

> A total of £9 million was allocated to Primary Care Trusts in 2010/11 to run local awareness and early diagnosis initiatives, targeting breast, bowel, and lung cancers.[840]
>
> Over three years, the Cancer Drugs Fund is making £200 million available annually to support improved access to cancer drugs.[841]
>
> Expenditure on cancer has increased from £3.19 billion in 2003/4 to £5.68 billion in 2012/13.This represents an increase of £680 million compared to 2011/12.[842]

These examples make it evident that the control over the purpose and the scope prioritises the variables to evaluate. The definition of the scope is political because …

> … Different parties or entities may have quite different objectives for programs or policy changes, and may apply different weights to various types of outcomes […] this can lead to the scope of studies being defined to address only one perspective, perhaps intentionally.[843]

Consequently, the identification of progress in the Cancer Strategy is narrowed to those topics selected by the stakeholder in charge. For instance, in the four progress reports published so far, the discussion of equality is only slightly addressed.[844] Instead, this has been a key topic of discussion in the public domain.[845]

A Strategy for Cancer. Fourth Annual Report." London, UK. P. 8.

[840] Department of Health. 2011e. "Improving Outcomes: A Strategy for Cancer. First Annual Report." London, UK. P.32.

[841] —. 2012b. "Improving Outcomes: A Strategy for Cancer. Second Annual Report 2012." London, UK. P. 58.

[842] Department of Health, Public Health England, and NHS England. 2014. "Improving Outcomes: A Strategy for Cancer. Fourth Annual Report." London, UK. P. 14.

[843] Hedrick, T. E. 1988. "The Interaction of Politics and Evaluation." *American Journal of Evaluation* 9(3):5-14. P. 7.

[844] The references to equality in the Cancer Strategy's reports are mostly limited to the work of the National Cancer Equality Initiative (NCEI) 'to support, inform and champion efforts to improve cancer outcomes by reducing inequality and promoting equality.' Department of Health. 2013d. "Improving Outcomes: A Strategy for Cancer. Third Annual Report." London, UK. P. 17.

[845] Some media in the UK have reported the existing inequalities in the country in the access to cancer treatment and the consequences of this 'postcode lottery'. See: BBC News. 2014b. "NHS's cancer help 'postcode lottery' angers Pembrokeshire nurse." London, UK, Campbell, D. 2014b.

The purpose and scope of the NAO study on the Cancer Strategy was the result of an internal process focused on VFM. Its aim was 'to reach a judgement on whether value for money has been achieved […] good value for money [is] the optimal use of resources to achieve the intended outcomes.'[846] This emphasis constrained the analysis to the identification of costs and benefits of the activities performed under the Cancer Strategy, which refers specifically to looking at its …

> … performance in delivering cancer services; on improving information on activity, cost and outcomes of cancer services; and how the Department intends to deliver cost-effective cancer services in the new NHS.[847]

As in an internal evaluation process, the NAO and the PAC control the definition of the purpose and scope. In the case of the NAO, this decision relies on the C&AG and senior management.[848] They have a formal space of discretion to decide those policy areas that will be analysed by the NAO.[849] The rationale behind these decisions is not very clear, but it can respond to multiple factors including the political relevance of a policy issue.[850] The focus on VFM clearly prioritises values like efficiency, economy, and effectiveness. As a MP (UK-12) recognises: 'If you want to save as many lives as possible, wouldn't it be better if you had more accurate and timely information about what's going on?' This is a political decision because the predomination of these values over others like equality or justice sends the message that the 'success' of a policy is understood as its capacity to use effectively its resources to deliver services, rather than perhaps solving or alleviating the problem. Under the light of politics, this control over the purpose and scope responds to the interests of stakeholders

---

"Cancer patients dying early because of postcode lottery " in *The Guardian*. London, UK.

[846] National Audit Office. "What is a Value for Money Study?". London, UK.

[847] Committee of Public Accounts. 2011. "Delivering the Cancer Reform Strategy." London, UK. P. 4.

[848] For further reference about the VFM process undertaken by the NAO see White, F. 1999. *Audit, Accountability and Government*. Oxford, UK: Clarendon Press.

[849] According to the NAO, in the development of the VFM programme, 'we take account of suggestions from the PAC, but the Comptroller and Auditor General alone decides on what the NAO will examine.' National Audit Office. "Planning the value for money programme." London, UK.

[850] The privatisation of the Royal Mail, for example, was subject of an important political debate and attention from the media. See: BBC News. 2014a. "Business Secretary Vince Cable defends Royal Mail sale." London, UK, National Audit Office. 2014c. "The Privatisation of Royal Mail." London, UK: National Audit Office, Post & Parcel. 2014. "NAO says UK lost hundreds of millions through Royal Mail IPO." London, UK.

who, as Mohan and Sullivan recognise, 'attempt to get evaluators to report to policymakers findings that will cast their programs in a favourable light.'[851] Therefore, since the establishment of its purpose, the stakeholder commissioning the evaluation defines what 'success' would look like.

This also affects the implementation of a policy and how outcomes are perceived by beneficiaries. The constraints imposed through the purpose and scope of a study (to efficiency and efficacy in this case) exclude from the discussion variables considered relevant by others. For example, in relation to cancer treatment, the Chief Executive of a cancer charity (UK-11) considers that better results for patients could be achieved if alternative mechanisms of attention would be considered: 'Personalised medicine should be able to avoid amounts of wastage, but the NHS is not designed for that sort of inefficiency.' However, if evaluation and audit instruments do not incorporate other kinds of variables, it will not be possible to obtain evidence that might support alternative options. This leads to questioning the purpose of evaluation in relation to the expected results for the commissioner (NAO, in this case). It appears that both the evaluation performed by the DoH and the VFM study by the NAO are directed to measure the progress of the operating policy and identifying possible courses of action to improve efficiency, but these instruments do not inquire about policy alternatives for addressing cancer or the pertinence of the current policy—to redesign, continuation, or termination.

The diversity of actions undertaken within the framework of the Cancer Strategy increases the difficulty of the assessment. The DoH provides an annual report on the implementation of the strategy and the behaviour of specific variables, e.g. mortality, incidence, and survival rates. This is also framed by the information that the DoH considers relevant and suitable for reporting progress, such as the best and worst performing Trusts from Cancer Patient Experience Survey and the analyses made by the NCIN.[852]

---

[851]Mohan, R., and K. Sullivan. 2006. "Managing the Politics of Evaluation to Achieve Impact." *New Directions for Evaluation* 2006(112):7-23. P. 9.

[852] Department of Health, Public Health England, and NHS England. 2014. "Improving Outcomes: A Strategy for Cancer. Fourth Annual Report." London, UK.

In synthesis, evaluation processes for the Cancer Strategy are mostly driven by two major purposes: the identification of VFM and the monitoring of key variables, e.g. mortality and incidence rates that can be associated to its operation. However, in both cases, the focus is on the values of efficiency, effectiveness, and economy.[853] This emphasis on VFM leads to discussing about how the meaning of 'success' of policy is built. When speaking about a sensitive topic like cancer, it appears that the predomination of the VFM approach is unable to cover the different areas that should be assessed. Nonetheless, its sensitiveness favours its permanency in the scrutiny agenda. As DoH official (UK-8) observes:

> Cancer is a very emotional subject, [...] every MP would have cancer patients in their constituency, there'd be a story, a heart-breaking story of a child or a person who's died of cancer, or somebody in their families, and so cancer's hold to high scrutiny by Parliament, so by the MP, so lots of parliamentary questions, lots of the parliamentary letters, lots of debates in the House, both houses, Commons and Lords, there's a high level of scrutiny of cancer by Parliament itself.

The political salience of cancer can also be observed in the attention given by the media and external stakeholders, whose perception and interests can be divergent as 'stakeholders frequently disagree on the nature and approach to the problem.'[854] For instance, the governmental response to cancer can be subject to different interpretations and opinions. In the context of this year's election, possible courses of action to treat this disease have been a topic of discussion among politicians.[855] As a very emotional subject, cancer generates public expectations about the way in which the government should respond.[856]

The most relevant lesson from this case regarding the purpose of evaluation is that the political decisions about which elements are worth analysing constrain

---

[853] The driver for these values within government can be observed in the premises established by *The Green Book* and *The Magenta Book* (see Section 4.1.2) in which impact assessment appears to be a predominant tool in the policymaking process.
[854] Fitzpatrick, J. L. 1989. "The Politics of Evaluation With Privatized Programs: Who Is the Audience?" *Evaluation Review* 13(6):563-78.. P. 564.
[855] See for example: BBC News. 2015. "Cancer care in England has lost momentum, say MPs." London, UK, Bostock, N. 2015. "Cancer care slipping under coalition government, MPs warn." London, UK.
[856] In the UK media for example it is possible to regularly find stories about people living with cancer and the difficulties of this process. See: AFP. 2015. "Oliver Sacks describes learning he has terminal cancer: 'My luck has run out' " in *The Telegraph*. London, UK.

it to those variables that are of interest to the stakeholder that controls this stage, i.e. the NAO, the DoH. According to Bjørnholt and Larsen, the establishment of goals allows politicians to make explicit their political priorities.[857] However, when looking at the establishment of an evaluation's purpose this could also be a mechanism for politicians to promote the alignment of political priorities, goals, and the outputs of an evaluation. This discussion finds echo in the relevance given to the institutionalisation of evaluation. In the absence of formal rules, the discretion of stakeholders guides the establishment of the purpose, taking into account the political salience of this issue in the public landscape.

Resources, i.e. data, staff, or funding, play a very important role in evaluation. As Bamberger suggests, its control is strictly linked to the political nature of evaluation.[858] In the case of cancer, there is a strong dependence on scientific evidence generated by all kinds of organisations.[859] Perhaps one of the most important characteristics of this case is the diversity of stakeholders engaged in this policy area willing to pressurise the government into pursuing their interests, and that might use evidence for this purpose. As the Chief Executive of a cancer charity acknowledges (UK-11), 'Cancer is a well-funded disease group compared to other disease groups, there's a lot of evidence in the cancer registries and there's an interest to use that data to make strategies.' Some other interviewees add: 'On cancer there's relatively quite strong information' (UK-4) and 'It's easier to make evidential statements rather than subjective statements which are harder to argue against to evidence-based cases' (UK-5)*.*

The objectivity associated with scientific evidence supports stakeholders' views about cancer and represents a valuable resource. For example, arguments that promote regulation changes in relation to smoking have been strongly based on

---

[857] Bjørnholt, B., and F. Larsen. 2014. "The politics of performance measurement: 'Evaluation use as mediator for politics'." *Evaluation* 20(4):400-11.
[858] Bamberger, M. 1991. "The Politics of Policy Evaluation in Developing Countries." *Evaluation and Program Planning* 14(4):325-39.
[859] In the governmental sector for example: NICE, NCIN and the DoH; in the social sector it outstands the work of NGOs such as Cancer Research UK and MacMillan Cancer Support; in the academic sector there are research centres and universities that systematically develop research about cancer, for example, the School of Health and Related Research at The University of Sheffield and the UCL Cancer Institute.

evidence.[860] However, the use and analysis of data can also be subject to politics, especially in its control and access, which are key elements for producing evidence and for the development of evaluations and audits.[861] According to the theory, this is an important area of political susceptibility.[862] Restraining or facilitating access can represent a way to influence the process. This leads to questioning the extent to which stakeholders involved in the production of cancer data are willing to share this information with others.

According to a recent PAC inquiry, the reform of the cancer networks did not make data access more efficient, as 'organisations, including cancer charities, have experienced delays in getting access to the data needed for research and analysis to support improvements in cancer services.'[863] The actor that controls data[864] can limit cooperation with the evaluation process.[865] Beyond this, the political issue to discuss is that even though there are other stakeholders who develop research in this field, they rely on government data for this purpose. As the NAO's report suggests, 'A number of organisations that we interviewed also highlighted that the flow of cancer data around the health system had reduced since 2013.'[866] Access to data as a resource for evaluation and auditing can become a political matter when its control depends on a sole actor who can intentionally or unintentionally block, delay, or deny access.

---

[860] As Cancer Research UK states 'smoking is by far the most important preventable cause of cancer in the world [...] 64,500 cancers are caused by smoking each year in the UK.'

[861] Denzin, N. K. 2009. "The elephant in the living room: or extending the conversation about the politics of evidence." *Qualitative Research* 9(2):139-60.

[862] Datta, L. E. 2011. "Politics and Evaluation: More Than Methodology." *American Journal of Evaluation* 32(2):273-94.

[863863] According to Sarah Woolnough, Executive Director of Policy and Information of Cancer Research UK, 'there has been a loss of national and local leadership and infrastructure resource as a consequence of the changes. There used to be, for example, a national cancer action team that helped Sir Mike deliver the cancer strategy—that was disbanded [...] the 28 cancer networks that existed have been reduced to 12 strategic clinical networks, and they are not cancer specific [...] the strategic clinical networks cover a range of disease areas.' Committee of Public Accounts. 2015c. "Oral evidence: Progress in improving cancer services and outcomes." London, UK: House of Commons.

[864] This has also been addressed by Bamberger et al. in relation to the relevance of the quality of data. For them, 'project records and other documentary data often suffer from reporting biases or poor record-keeping standards...' Bamberger, M., J. Rugh, and L. Mabry. 2012. *Real-World Evaluation: Working Under Budget, Time, Data, and Political Constraints*. California, US: SAGE. P. 25.

[865] Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106.

[866] National Audit Office. 2015b. "Progress in improving cancer services and outcomes in England." London, UK. P. 21.

Specialists available to assess the Cancer Strategy are another resource connected to politics. Going back to the idea of a market of knowledge it is clear that the use of external evaluators/consultants can provide more legitimacy to the evaluation process.[867] An official of the NAO (UK-3) recalls that some years ago 'there was a sort of a government drive to make use of consultants whereas now it has gone the other way.' This can be explained mostly because of the budget cuts that the public sector faces. According to *The Guardian*, for example, there are '70 projects, worth £8.9m over several years, that had been cancelled or cut short across four government departments since May last year.'[868]

Despite this trend, the role of academics in the development of policy studies is still important, particularly in a highly technical area like cancer. This is recognised by a DoH official (UK-8): 'You need academics to produce it with the quality standards of academics.' In this case, the prestige of academics gives cancer research more external legitimacy.[869] In the UK, there is an important range of evaluators, consultants, NGOs, and research centres developing high-quality cancer research.[870]

The hidden politics of evaluation are on the level of discretion in the commissioning of evaluations. Although this is based on a public bidding process with clear criteria for evaluating the proposals, an evaluator (UK-14) is sceptical about 'whether there is an unconscious bias in the selection [...] there might be sometimes when the commissioner already has in mind the right organisation to do the work but you have to put it out to tender anyway.' Discretion in the commissioning of evaluation was also analysed by the LSE GV314 Group[871] in terms of the extent to which the UK "government evaluation research tends to concentrate on 'safe' topics – that is, those less likely to

---

[867] Bogenschneider, K., and T. J. Corbett. 2010. *Evidence-Based Policymaking: Insights from Policy-Minded Researchers and Research-Minded Policymakers*. New York, US: Routledge.

[868] O'Hara, M., and J. Shepherd. 2011. "Government cuts millions from research and evaluation budgets " in *The Guardian*. London, UK.

[869] According to Conley-Tyler, 'the appearance of impartiality remains a real need among many evaluation audiences and remains a strong argument for the use of external evaluators.' Conley-Tyler, M. 2005. "A fundamental choice: Internal or external evaluation?" *Evaluation Journal of Australasia* 4(1 & 2):3-11. P. 8.

[870] For example, Cancer Research UK and McMillan Cancer Support.

[871] This group is 'a final year undergraduate course in the Government Department at LSE.' London School of Economics. 2014. "GV314 Empirical Research in Government."

produce results that will embarrass it."[872] According to this research:

> Government does appear in general to play it safe: 31 per cent of respondents agreed with the statement that 'Government is often prepared to take risks in commissioning research that might produce results that highlight shortcomings in policies and programmes', while 49 per cent disagreed and 20 per cent neither agreed nor disagreed (N = 185).[873]

This concerns also the development of evaluation capabilities. Despite there being no explicit purpose related to this, the Cancer Strategy shows that the use of social research has bloomed in this area, perhaps as a consequence of its political relevance—'It affects older people who vote and have time to write to their politicians' (UK-11)—and its impact on the UK's economy (£15.8bn.).[874]

Funding availability can determine, for example, the scope of a study, the consultant to hire, or the methodology to apply. This is explained by a NAO officer (UK-3): 'If we are doing a three-month study methodologies tend to be sort of simpler, interviews with stakeholders […] the Department, the various organisations, data analysis and probably not a lot else…' The constraints imposed by the resources at disposal play an important role in key decisions about the evaluation:

> Since 2010–11 four departments have reduced evaluation resources; four have cancelled or curtailed 25 evaluations between May and December 2010; eleven ongoing evaluations were cancelled before completion, reducing spending by more than £3 million; [and] a further 14 evaluations were cancelled. [875]

> 'We would then need to consider the methodology you'd propose in terms [of] how much that would cost, because sometimes short studies are not necessarily cheap, sometimes a three-month study can be as expensive as a 9-month study if the topic is big and would focus on a theme but that doesn't mean it would be cheap…' (UK-4).

The allocation of funds for evaluation or auditing can become a mechanism of political control.[876] As Chaplowe and Engo-Tjéga state, funding and sponsorship

---

[872] The LSE GV314 Group. 2014. "Evaluation under contract: Government pressure and the production of policy research." *Public Administration* 92(1):224-39. P. 228.

[873] Ibid. P. 228.

[874] This is the annual cost of all cancers to the UK economy according to an Oxford University study. University of Oxford. 2012. "Cancer costs the UK economy £15.8bn a year." Oxford, UK.

[875] National Audit Office. 2013a. "Evaluation in government." edited by National Audit Office. London, UK..

[876] Azzam, T., and B. Levine. 2014. "Negotiating Truth, Beauty, and Justice: A Politically Responsive Approach." *New Directions for Evaluation* 2014(142):57-70.

imply multiple challenges for evaluators, including having to make decisions about the scope or methodology in relation to the funding available.[877] It can also imply a compromising relationship with the funder.[878] In the case of the Cancer Strategy the safeguard is given by the institutional framework that clearly norms the public bidding process and the way in which public contracts must be allocated.[879]

The hidden politics of evaluation in relation to resources can be reduced to the establishment of constraints that might impact the outputs of an evaluation. Major restrictions in this case can be observed from the fact that the different actors involved, i.e. NAO and DoH, have limited resources to perform more sophisticated studies or to make a more systematic follow-up of evaluations. The continuity of the assessment is restrained not only by the political commitment of decision-makers, but also by the resources available and the need for prioritising evaluation areas. A member of the PAC (UK-12) realised that in terms of following up reports, 'We are not as good at it, we're not as systematic as we should be.' Therefore, following up evaluations also depends on the political agenda which might be linked to the fear of negative findings. For example, the discussion about the Cancer Drugs Fund[880] might have fostered the interest of the PAC in reviewing the government's response to cancer.[881] As discussed, the electoral climate has also brought more attention to this disease,[882] which can be translated into the negotiation of more resources

---

[877] Chaplowe, S. G., and R. B. Engo-Tjéga. 2007. "Civil Society Organizations and Evaluation: Lessons from Africa." *Evaluation* 13(2):257-74.

[878] Lipsky, D. B., R. L. Seeber, A. C. Avgar, and R. M. Scanza. 2007. "Managing the Politics of Evaluation: Lessons from the Evaluation of ADR Programs." Pp. 116-29 in *Proceedings of the Fift-Ninth Annual Meeting of the Labor and Employment Relations Association*.

[879] "The Public Contracts Regulations 2015."

[880] The Cancer Drugs Fund 'is money the Government has set aside to pay for cancer drugs that haven't been approved by NICE and aren't available within the NHS in England [...] the aim of the fund is to make it easier for people to get as much treatment as possible.' Source: Cancer Research UK. 2015b. "Cancer Drugs Fund." London, UK.

[881] Even though cancer has been a permanent subject of discussion in the public arena, after the NAO's report in 2011, it was until last year when the strategy was revised again.

[882] For instance, different NGOs in the country have made public manifestos in relation to cancer for the general elections of this year. See: Breast Cancer Care. 2015. "The Breast Cancer Manifesto 2015. Breast Cancer Care's ambitions for the next government." London, UK, Target Ovarian Cancer. 2014. "General Election 2015 Manifesto." London, UK.

for research about cancer.[883]

Perhaps it is in the *process* of evaluation that the influence of politics becomes more evident. The institutional framework in the UK promotes the openness of assessment processes through the inclusion of external evidence providers. In the NAO study and in the PAC's inquiry, different stakeholders were invited to give evidence.[884] A member of the Health Committee staff (UK-2) explains this in more detail:

> Sometimes we might put terms of reference if there are specific issues we want them to comment on [...] is announced through a press notice so that goes to the press and also to a lot other people on our list, it's put on our website as well.

Other interviewees agree in recognising that there are conditions for different actors to access evaluation processes and to openly put their views on the table about a specific policy issue:

> We have a good relationship with a lot of MPs as well because we do work with them, so I think we have good opportunity to put forward any concerns that we have… (UK-6)
> The Committee did feel it was very helpful to have outside interest giving the evidence before we have the official view… (UK-7)

The formal processes established for the participation of stakeholders suggest that the institutional framework operating in the UK favours a more inclusive approach. Considering the value given by the literature to this issue,[885] its political relevance demands acknowledging 'who key policy players are, their historical roles in framing policy debates, their current policy objectives, and the consequences of these policy players' influence on evaluation practice.'[886]

The process allows participation, but this *does not mean it is exempt from political influence.* Cancer represents a policy area with high levels of

---

[883] Balthasar, A. 2009. "Institutional Design and Utilization of Evaluation: A Contribution to a Theory of Evaluation Influence Based on Swiss Experience." *Evaluation Review* 33(3):226-56.

[884] The most recent PAC's inquiry included a range of governmental (i.e. NICE, NHS England) and non-governmental witnesses (i.e. Cancer Research UK and McMillan Cancer Support).

[885] See for example the work of Green regarding this point. Greene, J. C. 1997. "Evaluation as Advocacy." *American Journal of Evaluation* 18(1):25-35, —. 2001. "Dialogue in Evaluation: A Relational Perspective." *Evaluation* 7(2):181-87.

[886] Rutkowski, D., and J. Sparks. 2014. "The new scalar politics of evaluation: An emerging governance role for evaluation." Ibid.20(4):492-508. P. 504.

engagement. There are channels to take part in evaluation and audit processes. Nonetheless, there may be asymmetries between stakeholders. Larger organisations, i.e. CRUK, will have more resources for influencing the policy process than smaller ones.[887] Moreover, there are asymmetries of power and influence among different types of cancer groups. Regarding this, the Chief Executive of a cancer charity states (UK-11): 'I think Cancer Research UK should be more of a NGO and not a competitive charity […] but it has all the benefits of the NGO, so they have very strong access into government.' Through the lenses of the hidden politics of evaluation these asymmetries indicate that in addition to the initial allocation of resources of each stakeholder, there are political factors that favour or limit their participation such as the interest of politicians on a particular disease group.[888] This can be considered a bilateral relationship; politicians give more attention to a disease group because it affects a wider proportion of their electorate, and vice versa.[889]

Breast cancer, for instance, receives a great amount of attention from the government, and important efforts for the production of knowledge have been made in this particular disease group.[890] MPs from different parties have advocated breast cancer, giving it more notoriety than other disease groups,[891] generating reactions from other groups:

> There is an understandable but very frustrating bias towards the big cancers, so if you've got breast cancer there are many more different strategies and options available for you than if you get ovarian cancer,

---

[887] In terms of funding, CRUK reported having £490 million to carry out their activities, in contrast to, for example, Ovarian Cancer Action that reported in 2013 an income of £2.3 million. Sources: Cancer Research UK. 2014b. "Every Moment Counts. Our Annual Review 2013/14.". London, UK: Cancer Research UK, Ovarian Cancer Action. 2013. "Annual Report and Audited Accounts for the Year Ending 31st March." London, UK: Ovarian Cancer Action.

[888] Conservative and labour MPs, especially women, have made public their support to wide-known movements in favour of breast cancer victims, e. g. Pink October.. See: Mactaggart, F. MP. 2014. "Slough MP 'Wears it Pink' for Breast Cancer Awareness.", McKinnell, C., MP,. "Catherine goes pink for breast cancer awareness."

[889] Statistics show that breast cancer is the most common type of cancer in the UK. Source: Breast Cancer Care. "Facts and statistics."

[890] For example, the NICE Guidance on Cancer Services focuses on breast cancer; the DoH has also developed guidance for 'improving breast cancer clinical care' and the NCIN has piloted studies in relation to this type of cancer. In contrast to other disease groups, breast cancer has been a core element of the government policy. See: National Cancer Intelligence Network. 2012. "Recurrent and Metastatic Breast Cancer Data Collection Project. Pilot Report.".

[891] See: Moody, J. 2015. "Breast screening unit is now back where it belongs in Uttoxeter." in *Uttoxeter Adviser*.

> pancreatic cancer and I think that it's understandably complex for the health strategy to keep pace with cancer discovery. ... (UK-11)

Access and capacity of influence are linked to the resources and power of organisations and individuals. It is natural to observe that certain organisations become recurrent evidence providers and it is left to the discretion of the people in charge of the inquiry to incorporate or dismiss that evidence. For example, comparing the witnesses that took part in the most recent NAO study with the PAC's inquiry it is possible to identify that some evidence providers were dismissed in the latter and did not take part in the oral session.[892] There is an internal area of discretion in the selection and prioritisation of evidence:

> Mostly we receive evidence from the people and the organisations we always expect to receive evidence from, and we know what position they are coming from and what they are arguing in favour ... so we sort of know they are going to say that and so we can take that into account when we look at their views (UK-2).

> It's a judgement call to sort of how far down different groups you go, you have to decide who you think is important to talk to, whether you cover an issue or whether they sort add value that you need to speak to them because who haven't spoken to anybody who might have a view on that type of issue (UK-3).

Another important issue regards the relationship between stakeholders, which relates back to the idea of democratisation of evaluation.[893] According to Karlsson, this 'could be seen as an attempt to break the dominance of the one-sided commissioners' and contents of the evaluation power in favour of the aims and contents of the evaluation.'[894] The relationship between evaluated bodies and evaluators, for example, is an interesting element of analysis. In the case of the Cancer Strategy, this refers to the relationship between the DoH (as the implementation agency) and the NAO and PAC. In terms of the cooperation in the provision of information for the NAO's studies, some interviewees explain:

> Most of the time it's absolutely fine, I mean, under statutory law they're obliged to provide us with anything that we request [...] some

---

[892] Among them Breast Cancer Campaign and the King's Fund.
[893] Hanberger, A. 2006. "Evaluation of and for Democracy." *Evaluation* 12(1):17-37.
[894] Karlsson, O. 1996. "A Critical Dialogue in Evaluation: How Can the Interaction between Evaluation and Politics be Tackled?" Ibid.2(4):405-16. P. 405.

organisations are quicker than others at providing information, most of the time the relationship is fine, yet sometimes it can get particularly slow, so you might have to push it up the line so that somebody higher up in their organisation is informed that we are not getting the information that we require, but we do always get the information that we need (UK-3).

The NAO gets support from the PAC because when the PAC get a lot of publicity, which they do, it means that the NAO can go in, and they can go into the departments and they can increase the pressure on the departments to give them the information they need (UK-7).

A DoH official (UK-8) adds:

When the NAO produces a report we help them creating terms of reference [...] 'so they're not barking up the wrong tree', [...] so we help them making as useful as possible, they produce the evidence largely and gather the evidence.

This raises questions about the pertinence of keeping some distance between evaluators and evaluated bodies and how this affects independence and objectivity. According to Lonsdale's experience at the NAO:

There appear to have been increased informal discussions about study topics between the NAO and senior officials within audit bodies [...] Increased efforts have also been made to be responsive to departmental views on the methods (their appropriateness, for example), the timing of work (from specific visits to the study as a whole) and the consultants used (to avoid conflicts of interest and ensure their credibility with those under examination).[895]

The recognition of this 'closeness' gives important insight into the hidden politics of evaluation.[896] In the Cancer Strategy, for example, the fact that the DoH participates in the definition of the Terms of Reference (ToR) that will guide the evaluation might interfere with the objectiveness and independence of the inquiry. It may not be a deliberate attempt to interfere in the NAO's work; however, it can become a mechanism to influence the scope, purpose, and outputs of a study. This brings back House's ideas about the appropriate distance that must be kept between stakeholders.[897] On the one hand, a more

---

[895] Lonsdale, J. 2008. "Balancing Independence and Responsiveness: A Practitioner Perspective on the Relationships Shaping Performance Audit." Ibid.14(2):227-48. P. 235.

[896] Lipsky, D. B., R. L. Seeber, A. C. Avgar, and R. M. Scanza. 2007. "Managing the Politics of Evaluation: Lessons from the Evaluation of ADR Programs." Pp. 116-29 in *Proceedings of the Fift-Ninth Annual Meeting of the Labor and Employment Relations Association*.

[897] House, E. R. 2008. "Blowback: Consequences of Evaluation for Evaluation." *American Journal of Evaluation* 29(4):416-26.

operational perspective on the inquiry might provide guidance about specific and technical issues.[898] On the other hand, too much closeness between actors can lead to a biased judgement.[899]

The political implication to debate here is the degree of influence that government departments have in the outputs of a study. Specifically in the case of VFM studies (see Section 4.1.3), this can be perceived at the 'clearance' stage in which the NAO sends report drafts to government departments 'to check that they're happy with the factual accuracy' (UK-3).[900] A DoH official explains (UK-8): 'We see drafts, we have the right to comment on drafts, I don't think we have the right to change.' However, it seems interesting that during the interviews the idea of reports based on 'agreed evidence' was recurrent. This means that the data used for the studies is derived from a process in which both the auditor and the audited body work collaboratively to build the evidence. This can lead to political struggles in which the final outcome will be driven by the interests of the stakeholder who exercised its power more strongly. Regarding this Sharma observes that:

> 'Clearance' is a convention whereby the facts and presentation of VFM reports are 'agreed' with Accounting Officers before publication (NAO, 2003a). The formal process involves two key stages of clearance, between the audit team and the Principal Finance Officer and the main representatives of the Agency under audit.[901]

It was denied by interviewees that the outputs of a study can be manipulated during 'clearance'. Still, when speaking of a possible negotiation between the NAO and government departments, an interviewee (UK-3) explains:

---

[898] The work of Owen et al. about the evaluation of the Early Excellence Initiative provides an interesting discussion about the interaction between stakeholders in the process of evaluation, including practitioners. Owen, J., T. Cook, and E. Jones. 2005. "Evaluating the Early Excellence Initiative: The Relationship between Evaluation, Performance Management and Practitioner Participation." *Evaluation* 11(3):331-49.

[899] See: House, E. R. 1997. "Evaluation in the Government Marketplace." *American Journal of Evaluation* 18(1):37-48.

[900] The clearance stage found in the NAO's VFM process has been questioned by different scholars in the audit field. See: Gray, A. , and B. Jenkins. 1993. "Separate Developments: Budgeting, Auditing, and Evaluation in the United Kingdom." in *Budgeting, Auditing, and Evaluation. Functions and Integration in Seven Governments* edited by A. Gray, B. Jenkins, and B. Segsworth. London, UK: Transaction Publishers, Sharma, N. 2007. "Interactions and Interrogations: Negotiating and Performing Value for Money Reports." *Financial Accountability & Management* 23(3):289-311.

[901] —. 2007. "Interactions and Interrogations: Negotiating and Performing Value for Money Reports." *Financial Accountability & Management* 23(3):289-311. P. 298.

> I think negotiation is possibly not the right word [...] in some ways it is the right the word, I mean they'd put forward arguments if they really don't want a study to be done at a certain time, but as I said, it's very rare, I mean, the most that would normally happen is that the study would get put back.

At least informally, there are spaces of discretion in which political power can be exercised, particularly in those cases where actors hold an important amount of power, i.e. C&AG:

> At the end of the day it is the C&AG's decision if he wants to do a study and even if there are objections, he might say this issue is too important we need to do a study now and he would drive that through ...(UK-3).

A final remark concerns the evolution of the relationship between the NAO and the PAC. As the NAO represents a technical body that feeds into the accountability processes undertaken by the PAC, it would appear that the interests of these two institutions might be aligned, but according to a PAC's former clerk (UK-7):

> The relationship has changed in the last four years, in previous times the NAO would put forward a draft report to the PAC and they would always accept it, almost always accept it except for one or two words, now there's often a process of discussion in the Committee about whether the wording of the report is right and this occasionally take a party political tension.

This statement reveals political discrepancies between them showing that conciliating interests does not occur *de facto*. The strength of the PAC in the House of Commons inevitably can generate political disputes. As a member of this Committee (UK-12) explains: 'Our committee works completely differently, instead of doing [like] most committees would do sort of 7, 8, to 10 reports a year, maybe we do 50, or 60.' This amount of inquiries covering a wide range of policy areas would lead to the confrontation of different views, not only between the PAC and the government, but also with the NAO and other stakeholders. The discussions about the Cancer Drugs Fund and the performance of NICE in this area illustrate the kind of disagreements that can arise.[902] In this case, for instance, CRUK believes ...

---

[902] NICE plays a key role in relation to the assessment of those drugs that should be funded by government and this year is taking place a very active debate in relation to the VFM of the CDF According to Professor Karl Claxton, of York University: 'Last year the Cancer Drugs Fund spent £280m on medicines [...] that money did some good but it would have done a lot more if it had

> ... that NICE should be reformed to ensure sufficient flexibility in the assessment of cancer drugs so that a separate Fund would no longer be considered necessary [...] proposals to reform NICE (Value Based Assessment) were put forward but rejected by stakeholders in 2014 as they were unlikely to significantly change NICE's approach to assessing new drugs, and may have disadvantaged some patient groups.[903]

The government defended its position by stating that:

> The decisions on the cancer drugs fund were not just to look at what drugs may not be supported, but also to be mindful that there are newer and better drugs available to us as a result of the challenge that we gave the pharmaceutical industry to provide more personalised drugs for different clinical indications [...] our objective is to be able to get the best drugs for the best value for money in the cancer drugs fund.[904]

These debates are useful for identifying more clearly the role of politics in evaluation and auditing. Despite the evidence at disposal and the criteria used for the evaluation, discussions also emerge from the discrepancies between the values of stakeholders. While for the government a better allocation of funds is considered a priority, disease groups consider that this approach does not capture the weight of other important variables, i.e. justice, equity. Stakeholders give different weights to evidence for building their arguments, prioritising those elements that support their positions. The Cancer Drugs Fund is a very clear example of how politics can overtake evidence. Technically, there is little to argue against the fact that this fund is inefficient, and yet despite the evidence, the discussion becomes political in terms of which values and interests should predominate: efficiency or equity?[905] It also puts the light on

---

been spent elsewhere in the National Health Service.' McKie, R. 2015. "David Cameron's flagship Cancer Drugs Fund 'is a waste of NHS cash' " in *The Guardian*. London, UK.

While the government aims to cut back funding for this purpose; other stakeholders argue that this could represent a serious problem in terms of equality for lots of patients in the country.("Mark Flannagan, chief executive of Beating Bowel Cancer, said two-thirds of patients with advanced bowel cancer 'face the probability of an earlier death' by being refused the treatments [...] 'Doctors will be forced to tell their patients there are treatments that can prolong their lives but will no longer be available', he said." Cooper, C. 2015. "Cancer Drugs Fund: Life-extending drugs to be denied to NHS patients in England as fund overspends " in *The Independent*. London, UK.

[903]Cancer Research UK. 2015f. "Written evidence from Cancer Research UK." London, UK.

[904] Statement made by Sean Duffy, National Clinical Director for Cancer, during the oral evidence session held by the PAC on the 21st January 2015. See: Committee of Public Accounts. 2015c. "Oral evidence: Progress in improving cancer services and outcomes." London, UK: House of Commons.

[905]Appleby, J. 2014. "The Cancer Drugs Fund: inequitable and inefficient?". London, UK: The King's Fund

the agenda of other stakeholders, e.g. pharmaceuticals,[906] which relates to one of the most relevant criticisms about the policymaking response to cancer. Most of the efforts have focused on treatment rather than on prevention:

> Researchers and their sponsors focused their efforts, and clinicians and their patients centered their hope more on the eradication of advances cancer than on its prevention or detection in surgically curable early stages.[907]

This suggests that cancer, both as a government policy and as a business area, represents important interests for stakeholders such as pharmaceuticals and medical infrastructure enterprises.[908] Despite the political confrontation about the government's response to cancer, the point to highlight here is that this has neither affected the credibility of evaluation nor the evidence produced about cancer. There is a debate based on different perspectives about how this disease should be addressed, but this does not affect the legitimacy of evaluation.

Outputs are another variable in which politics manifests through the use and dissemination of findings. The results from an evaluation process can have different uses, most of them already defined at the planning stage.[909] The Cancer Strategy case shows that the work of the DoH in terms of evaluation is more related to knowledge generation than to setting a judgement about this policy.[910] Outputs of this process appear to be used internally for decision-making.[911]

---

[906] See: Pharmaphorum. 2015. "Pharma 'shocked' as drugs cut from Cancer Drugs Fund."

[907] Faguet, G. B. 2008. *The war on cancer: An anatomy of failure, a blueprint for the future*. London, UK: Springer. P. 146.

[908] In addition, there has been research about the lack of profit of prevention efforts and the predominance of the interests of the health industry. See: Read, C. 1995. *Preventing breast cancer: The politics of an epidemic*. London, UK: Pandora.

[909] Bogenschneider, K., and T. J. Corbett. 2010. *Evidence-Based Policymaking: Insights from Policy-Minded Researchers and Research-Minded Policymakers*. New York, US: Routledge.

[910] As it was discussed before in this Chapter, the label of 'strategy' is a tool to integrate a series of activities and processes carried out by the government in the attention of the cancer problem, this considerably increases the complexity of evaluation because it is difficult to assess it as an integrated unit, instead, assessment processes have attempted to analyse and understand different initiatives related to this disease.

[911] For example, the implementation of standardised tobacco packaging was a decision taken based on different studies performed and commissioned by the DoH. Chantler, C. 2014. "Standardised packaging of tobacco." London, UK: King's College London, Department of Health, Public Health England, and NHS England. 2014. "Improving Outcomes: A Strategy for Cancer. Fourth Annual Report." London, UK.

External audit processes have a different use. These reports serve as a basis for inquiries in which those responsible for the operation of the strategy attend oral hearings at the House of Commons to debate the findings[912] and where the government responds to PAC's recommendations.[913] Saunders clearly identifies the political nature of evaluation outputs:

> The organizational environment into which an evaluation output might be intervening and the design of the evaluation output itself [...] interact to determine the extent to which an evaluation output or an evaluation process can create effects (i.e. change practices).[914]

Considering this, the context in which the NAO operates provide more leverage to audits than to evaluation, i.e. follow-up by the PAC. Nonetheless, the use of the findings is left at the discretion of stakeholders. An interesting finding about the Cancer Strategy case concerns the response of the DoH at the potential publication of the NAO's report:

> It was kind of [a] coincidence that as we've only started the fieldwork, shortly afterwards the Department decided to refresh the second strategy, so they published the strategy roughly about a month after our report, so lots of dialogue with commissioners in the field-work already fed into their new strategy (UK-4).

The government was aware of the possible areas of improvement to be reported by the NAO and they anticipated the reform of the strategy by incorporating several of the observations made in the study. This reflects that the evaluation made by the DoH and the audits performed by the NAO serve different purposes. On the one hand, the lack of formal mechanisms to foster the use of evaluation outputs confines them mostly to support decision-making within the DoH.[915] On the other hand, the close linkage between the NAO and the PAC

---

[912] The former National Cancer Director, Sir Mike Richards, attended an oral hearing with the PAC at the House of Commons, Mr David Flory, Deputy Chief Executive and Director General Finance and Christine Connelly, Chief Information Officer, from the Department of Health also participated in the session. Source: Oral Evidence Taken before the Committee of Public Accounts on Tuesday 7 December 2010.

[913] HM Treasury. 2012a. "Treasury Minutes. Progress on implementing recommendations from 28 Committee of Public Accounts reports (Session 2010-12) and 5 National Audit Office reports." London, UK.

[914] Saunders, M. 2012. "The use and usability of evaluation outputs: A social practice approach." *Evaluation* 18(4):421-36. P. 428.

[915] According to several interviewees, government departments are not obliged to follow recommendations, however, they are expected to respond to a Select Committee Report within eight weeks from the publication. In the case of the Cancer Strategy, the government through the

allows audit work to comply its accountability function in relation to how money is spent.

In terms of dissemination, there is no formal mechanism to enforce government departments to make evaluation reports public. This affects the process at different levels. The first one regards transparency, about which areas are being evaluated; for what reasons and by whom.[916] The second concerns the way in which evaluation findings are made public. According to the theory, 'political entities may disseminate or use evaluation results selectively.'[917] With no formal guidelines to enforce dissemination this becomes another area of discretion for stakeholders, especially commissioners. These pressures can be considered political because of the potential harm to the interests of stakeholders and the decisions they made for disseminating findings. A cancer researcher (UK-5) emphasises that there is no intentional lack of transparency, although he recognises that when it is a sensitive topic ...

> ... there are a lot of pressures to, I suppose, not make waves in a sense, but that's usually a case of publishing the data quietly rather than putting it in a big press release, making a big splash of it.

A policy think tank member (UK-18) adds that occasionally departments release evaluation reports on dates when they are not very likely to get attention, e.g. holidays.[918] As evaluation serves legitimisation purposes, evidence can be used to argue in favour or against a policy. It 'has become a tool for control and measuring to answer if the performances can show quality and efficiency and to legitimize the political decisions and priorities.'[919] Despite cancer being an evidence-based policy, a DoH official (UK-8) recognises that:

Treasury Minutes reported in 2012 the progress on the implementation of recommendations made by the PAC. Committee of Public Accounts. 2013a. "Select Committees. Government Responses." London, UK.

[916] As the NAO observed: 'It is difficult to establish the coverage of evaluation evidence, but it does not appear to be comprehensive.' National Audit Office. 2013a. "Evaluation in government." edited by National Audit Office. London, UK. P. 7.

[917] Hedrick, T. E. 1988. "The Interaction of Politics and Evaluation." *American Journal of Evaluation* 9(3):5-14. P. 8.

[918] An interesting example of this was the release of the report *Applying Behavioural Insights to Organ Donation: preliminary results from a randomised controlled trial*, published by the DoH last year on Christmas Eve.

[919] Karlsson, O. 2003. "Evaluation Politics in Europe: Trends and Tendencies." *Studies in Educational Policy and Educational Philosophy*. P. 1.

> The strategy is essentially political, it's decided by government ministers, so the government of the day is the strategy, so when's the election next May the government changes or may not change, that would be the lead before the strategy (UK-8).

The elections have fostered the engagement of politicians—especially from the opposition—in the debate about the performance of the Coalition Government. In the context of the PAC's report on cancer, a Labour MP expressed:

> This report makes devastating reading for David Cameron and is damning proof of his mismanagement of the NHS. Under the Tories, cancer care has gone backwards and patients are missing out on the treatment they need.[920]

This statement reflects a clear political use of evaluation/audit outputs. The use of evidence considered legitimate to emphasise negative findings is an attempt to politically discredit the party in office and to make explicit political promises in this area.[921] In line with this, a DoH official (UK-8) explains:

> We will find out when the political parties produce their manifestos they may well say something on cancer, then we'd have to start planning for that, as a department of state you plan for all eventualities, so if there's something on the manifesto of the Labour Party or the Conservative Party or even the Democratic Party, we have to make sure that what they're saying, if they get into power, how would we do that, all the preferences go prior general election.

Healthcare is a constant topic in the manifestos of political parties. For the recent election, for instance, major parties made explicit their intended actions in this area. The Labour Party 'will make sure the NHS is focused not on competition but on collaboration, so that the system is properly integrated to work in the interests of patients,[922] while the Liberal Democrats will 'invest in research and set ambitious goals to improve outcomes for the most serious life-threatening diseases like cancer.'[923] The use of evaluation outputs as political ammunition to discredit the work of the current government and to promote a partisan agenda is a common practice (especially during electoral times) and it

---

[920] Labour Press. 2015. "Response to PAC report on deteriorating cancer services."
[921] For example: Brady, E. 2014. "Labour Pledges One-Week Cancer Diagnosis." London, UK: Sky News.
[922] Labour Party. 2014. "Patients before profits."
[923] Liberal Democrat Party. 2014. "Liberal Democrat Pre-Manifesto 2014."

also draws more attention to evaluation and audit reports.[924]

The last variables to examine are the *outcomes* of evaluation. The long-term effects of assessment processes in policymaking are very difficult to observe and this has been one of the most relevant areas of concern in the evaluation studies field.[925] It becomes really complex to establish causal relationships between changes in a particular policy and the findings of an evaluation, particularly in those areas that involve variables that cannot be quantified. As Selameab and Stuart state: 'Social outcomes […] do not always lend themselves to expression in monetary terms.'[926]

In the case of a multifactorial problem like cancer it is almost impossible to isolate the net contribution of evaluation to its improvement. Positive cancer outcomes can be explained by numerous reasons, i.e. changes in lifestyle, medical history, not only as the result of government's intervention. As a researcher recognises (UK-5), 'There's no counterfactual, so if you succeed or if you fail was that because of the strategy or was that because of other facts beyond the strategy?' The establishment of indicators of progress, as those established by the DoH, i.e. mortality, survival rates, does not provide a complete panorama of the situation. Proctor explains this more clearly:

> Survival rates are only one of several ways to measure progress against a disease. Survival rates will not tell you if chemotherapy or radiation is less unpleasant than it was once, or whether surgery leaves the typical cancer patient less disfigured that it once did […] or whether patients are less likely to die from a cancer caused by medical treatment.[927]

Another element to consider is the level of integration of evaluation and audit mechanisms in policymaking. While in the past the utility of evaluation was strictly related to changes in policy, 'More recently, investigators have

---

[924] The recent NAO's report on cancer and the subsequent PAC's report generated circa 100 articles during March 2015. Source: Google UK.

[925] See Hughes, M., and T. Traynor. 2000. "Reconciling Process and Outcome in Evaluating Community Initiatives." *Evaluation* 6(1):37-49, Lloyd, N., and L. Harrington. 2012. "The challenges to effective outcome evaluation of a national, multi-agency initiative: The experience of Sure Start." Ibid.18:93-109.

[926] Selameab, T., and S. S. Yeh. 2008. "Evaluating Intangible Outcomes: Using Multiattribute Utility Analysis to Compare the Benefits and Costs of Social Programs." *American Journal of Evaluation* 29(3):301-16. P. 302.

[927] Proctor, R. 1995. *Cancer wars: How Politics Shapes What We Know And Don't Know About Cancer*. New York, US.: BasicBooks. P. 252.

considered the possibility that evaluation research is used in ways other than in directly affecting policy.'[928] From an instrumental perspective, the audit to the Cancer Strategy produced important improvements in the data-collection area (UK-4), also as a cancer researcher (UK-5) recognises:

> The data I provided for that audit on cancer [the NAO study] staging might have a huge political impact on the way cancer registration was performed, although I wouldn't claim it directly to what ended up happening in the subsequent years.

It would be rather simplistic to establish a direct link between the assessment of the Cancer Strategy and the changes that occurred after the publication of the reports, as these can have been driven by many other motives (including political momentum). Nonetheless, the way in which outputs are managed can provide some insights about the long-term effects of evaluation and audit. As discussed, the use of evaluation outputs derived from the DoH and other governmental bodies is mostly limited to internal consumption; *ergo,* little effect can be expected outside this sphere. Regarding the products derived from the NAO and the PAC, as it implies more public notoriety, they appear to have more effect in the public landscape, for example, by feeding into the discussions between politicians, government officials, media, NGOs, and academicians.[929] Finally, the effect that might be produced as a consequence of evaluation/auditing can lead to a dispute of power in terms of who is responsible for a particular change; in other words, who gets recognition. In the words of a member of the NAO (UK-3)*:*

> The department may not always want to admit that it has made those changes because we've done the study, but it was a driver, it would just never be admitted, so you're creating change almost just by sort of focusing attention on the area.

The Cancer Strategy shows that although it is a policy issue that heavily relies on evidence, it is not exempt from political influence. The asymmetry of access of stakeholders to evaluation processes affects the way in which political power can be exercised. Those stakeholders with more resources can influence key

[928] Boyer, J. F., and L. I. Langbein. 1991. "Factors Influencing the Use of Health Evaluation Research in Congress." *Evaluation Review* 15(5):507-32. P. 509.
[929] See: Campbell, D. 2015. "Government accused of breaking promises on NHS cancer treatment " in *The Guardian*. London, UK.

elements of the evaluation process like funding, methodology, or scope of an inquiry. In the broader context of governance, this case shows the linkage to accountability because the information created by evaluation and auditing serves as a basis for questioning the government's accomplishments and to debate the possible amendments that could be made in this area. The next case focuses on the CQC, an organisation that has been since its inception surrounded by controversy and, consequently, involved in a more subjective assessment.

6.1.3 The Hidden Politics of Evaluation in the Care-Quality Commission

The case of the CQC represents an interesting object of analysis because of its evolution and notoriety in the public landscape. Understanding the rationale behind the decision of merging organisations into a super-regulator of health and mental care is one of the most important elements in the context of the politics of evaluation. The establishment of the CQC was driven by political purposes, related to pressures for more efficiency in this field. Nonetheless, this decision provoked important criticism:

> A lot of the changes in the public sector in England have really been driven by the 2008 recession where we looked to reduce spend on public service provision, and I think the pressure to merge regulators therefore, get better efficiencies, better economies of scale have been driven politically (UK-1).

> It was created not from any obvious policy logic but simply from a desire to save money [...] as CQC came into existence it was the third new quality regulator for health, and the fourth for social care, in only nine years.[930]

> Its creation contained four big mistakes. (i) Its budget was inadequate for the purposes [...] (ii) It was rushed and was a forced creation with no visible attempt to learn from evidence in business mergers about the problems that always result in the short term from such imposed changes (iii) The rhetoric around its creation was about developing a tough, no-nonsense approach – so the people recruited to senior jobs displayed those characteristics – and we have just seen the consequences of that.[931]

The controversy around the CQC's inception has played a crucial role in the discussion about its effectiveness. Moreover, past events that involved the CQC,

---

[930] Timmins, N. 2013. "Designed to fail: the Care Quality Commission is a warning to quango cullers." London, UK: Institute for Government.
[931] Greig, R. 2013. "CQC – 3 Big Reasons It Went Wrong; 3 Big Actions to Get It Right." London, UK: NDTi National Development Team for inclusion.

i.e. Morecambe Bay, Mid-Staffordshire, have also influenced the process of evaluation and auditing to this organisation. Consistently with the previous case, the analysis of the hidden politics of evaluation in the CQC will be guided by the thematic framework proposed in this research (Figure 6.4).

Figure 6.4 Overview of the Hidden Politics of Evaluation in the CQC

Care Quality Commision	**Purpose**
	Focus on VFM and the identification if the CQC is 'fit for purpose'
	Predefined by audit bodies (NAO, PAC, Health Committee) *Need for legitimisation due to negative public profile*
	Resources *Discretion in the selection of evidence providers* *Lack of transparency about evaluation costs*
	Process
	Distortions generated by relationship with DoH *Potential conflict of interests with evidence providers* *Influence of external stakeholders*
	Outputs
	Difference between accepting and implementing recommendations *Political momentum favoured use of findings*
	Outcomes
	Difficulty associated to measuring the effect of evaluation in the organisation *Need for better public profile (regain public trust)*

The establishment of the *purpose* of evaluation in this case has also been strongly driven by the values of efficiency, effectiveness, and economy. The DoH as the organisation in charge of supervising the work of the CQC has the task to evaluate its performance:

> To provide robust assurance to the public, the Department and Parliament that CQC is improving its performance and that action will be taken to build and sustain its capability for the future.[932]

The DoH review was based on three major sources of information: performance data, interviews with the CQC and stakeholders, and documentary evidence. Its aim was identifying those challenges that the organisation faced in its operation, while the focus was on the accomplishment of quantifiable measures, i.e. capacity to registry providers and inspections undertaken. This instrument

---

[932] Department of Health. 2012c. "Performance and Capability Review. Care Quality Commission." London, UK. P. 5.

allows looking at those areas of improvement in terms of the values underpinning it (efficiency, efficacy, effectiveness). As the theory suggests, the purpose reflects the role given to evaluation in the context of governance.[933] Considering that the creation of the CQC was driven by efficiency and efficacy values, it is logical to find these values behind evaluation. There is a need to produce evidence about the improvements made in efficiency that can legitimise the decision of its creation.

These values narrow the evaluation scope mostly to operational variables, leaving aside important elements related to this area, such as patient dignity and quality of care which would imply evaluating variables such as accessibility, focus on patients, equity, and safety.[934] However, the stakeholder that controls evaluation (the DoH, in this case) has the prerogative to define the purpose and the values associated to it, even when other stakeholders might have a different opinion about what must be the focus of evaluation, as a researcher of a social-care charity (UK-6) suggests:

> For me the best way to assess is what patients are saying, what people are saying, are people getting the outcomes that they need? Are people able to live life in a way that gives them independence?

The politics related to this stage concerns the establishment of rights and obligations in the institutional framework. For instance, the agreement between the DoH and the CQC states that the former's obligations include:

> Monitoring the whole system's performance to ensure it delivers what patients, people who use services and the wider public value most.[935]

Even without a formal evaluation policy, the institutional framework distributes powers among stakeholders that allow controlling the purpose and scope of evaluation, including decisions about when to evaluate. For instance, since the CQC began operation, the DoH has only formally reviewed its performance once (in 2012) in the context of the strong criticism made to this regulator as a result

---

[933] Berriet-Solliec, M., P. Labarthe, and C. Laurent. 2014. "Goals of evaluation and types of evidence." *Evaluation* 20(2):195-213.

[934] World Health Organisation. 2006. "Quality of care: A process for making strategic choices in health systems.". France.

[935] Care Quality Commission. 2013d. "Framework Agreement between the Department of Health and Care Quality Commission." London, UK. P. 8.

of whistle-blowing events. The lack of formal rules to define periodicity and time frameworks for evaluation gives the stakeholder in charge the capacity to decide this discretionally.

In the case of NAO study and the follow-up inquiries of the Health Committee and the PAC, the purpose was also to obtain evidence about the CQC's performance and its capacity to fit the purpose of its creation.[936] Considering the shortcomings observed in the CQC's performance (see Chapter 4, Section 4.3.2), the Chair of the Health Committee recognised that it …

> … is an acknowledged fact—and you have both acknowledged it in public many times—that the CQC, until recently, has not been delivering the service that the public looked to it to deliver.[937]

The focus of the inquiry was also driven by the capacity of the CQC to properly regulate and inspect health services. It becomes relevant to have a clearer perception of how these bodies define and prioritise the areas of assessment. The definition of the purpose and scope can be driven by multiple reasons, not all of them political. Financial and time constraints can lead to reducing the scope of an evaluation, while time frameworks can also affect its broadness.[938] In the case of the CQC, as a member of the Health Committee staff (UK-2) acknowledges, the purpose and scope of the assessment …

> … is driven by the members and how they evaluate the Care Quality Commission through the feedback that they get from members of the public, particularly of their constituencies, and from meetings that they have continually with different parties who are affected by the work of the CQC.

This reveals that the participation of MPs in scrutiny activities is embedded in a political process that inevitably affects their task. Most of their work would be shaped by the interests of members of their constituencies or by interest groups around them.[939] The inclusion of an issue in the PAC's agenda depends on the

---

[936] See: Health Committee. 2011a. "Accountability hearing with the Care Quality Commission." London, UK: House of Commons.
[937] —. 2014a. "2013 Accountability Hearing with the Care Quality Commission." London, UK.
[938] Mohan, R., and K. Sullivan. 2006. "Managing the Politics of Evaluation to Achieve Impact." *New Directions for Evaluation* 2006(112):7-23.
[939] For example, Andrew George MP, member of the Health Committee has expressed in relation to the work of the CQC in nursing homes of his constituency. See: Andrew G. MP. 2013. "Welcome for Cornwallis Care Quality report."

role of MPs. In addition, the institutional framework gives the PAC the prerogative to influence the work agenda of the NAO by proposing/suggesting potential studies. As a member of the PAC (UK-12) explains, 'the fact that a member of the Committee is very interested might help push it', so a policy issue that gets the attention of a MP is more likely to be considered in the NAO's agenda. This interviewee recalled the case of a PFI hospital[940] in the northern part of the country:

> The NAO did a report on that and I know it was partly because I was involved locally and I was a local member of Parliament, and they took a closer interest because I kept on going on and on about it (UK-12).

A similar case was the Leytonstone Hospital, in which the CQC reported 'poor standards of care and a culture of staff bullying.'[941] This situation was highlighted by a Labour MP, giving more leverage to this report's findings.[942] As a technical body of the PAC, the NAO sets the basis for most of this Committee's work. As a former clerk of this Committee describes (UK-7), the work of the PAC is driven by the inquiries of the NAO:

> The PAC is the public face of that process, and the criteria are these reasonably large amount of money that's wasted, is it of strong political interest, and thirdly, does it help the NAO to do its job of ensuring that public money is being well-spent, whether the committee can add value with an oral evidence session ...

The NAO also influences the way a study is designed and here the political preferences and interests of the members of the Committee play an important part in this process. This can increase the pressure to scrutinise the performance of a public body, as it occurred in the CQC case. An NAO official (UK-4) recalls that 'the CQC received a lot of media coverage before we started', so the need to study in detail the performance of the CQC became more urgent. The public profile of the CQC also influenced how its assessment developed: 'There were high political expectations for the newly created body' (UK-10).

---

[940] A Private Finance Initiative (PFI) allows 'enabled local authorities to enter into contracts with the private sector for the provision of new and/or improved capital assets (infrastructure for example) and related services.' Source: Department for Communities and Local Government. 2015a. "2010 to 2015 Government Policy: House Building." London, UK.

[941] O'Brien, Z. 2015. "Whipps Cross Hospital has been placed in special measures by the Care Quality Commission over bullying." in *Guardian*.

[942] Care Quality Commission. 2014c. "Whipps Cross University Hospital. Quality Report." London, UK.

Consequently, its evaluation and auditing were also shaped by these expectations.

The Winterbourne View and Mid-Staffordshire events brought important attention to the work of the CQC, which could have been reflected in more external pressure for scrutiny bodies to intervene. The NAO reported on the CQC's performance on December 2011[943] as a consequence of a petition made by the PAC. This was 'a quick reactive study' (UK-3) followed by an accountability hearing led by the Health Committee in September 2011[944] and a PAC's inquiry in May 2012.[945] In addition, the CQC faced important internal struggles as a result of the discrepancies between Board members that included questioning the credibility of the former Chief Executive.[946] During its first years of operation, the CQC 'wasn't working; people were falling through the gaps and nobody would take responsibility … there was no consistent standard' (UK-6). This was undeniably an opportunity for politicians to be proactive in addressing a problem that affects the general population and that had an important level of notoriety. In other words, to get advantage of this impetus for evaluating the CQC.

Likewise, a former clerk of the PAC (UK-7) explained that political interest does play an important role in the definition of the issues discussed within this Committee: 'It would have to be a wide-spread political interest' and the case of the CQC clearly showed that the failings of this regulator needed to be scrutinised. The inquiry made by the Health Committee in 2011, acknowledged:

---

[943] National Audit Office. 2011a. "The Care Quality Commission: Regulating the quality and safety of health and adult social care." London, UK.

[944] Health Committee. 2011c. "Annual accountability hearing with the Care Quality Commission." London, UK.

[945] Committee of Public Accounts. 2012. "The Care Quality Commission: Regulating the quality and safety of health and adult social care." London, UK.

[946] The former Chief Executive of the CQC, Cynthia Bower, was strongly criticised for her performance as Chief Executive of the West Midlands Strategic Health Authority (SHA) and for being involved in the Mid Staffordshire NHS Trust where 'as many as 1,200 patients may have died due to poor care.' Borland, S. 2012. "Care quango chief is forced to quit over a catalogue of failures (but she walks away with a £1.35m pension pot)." in *The Daily Mail*. London, UK.

In addition, she was also accused 'of conspiring in a cover-up over the organisation's failures to protect the public.' Sawer, P., and L. Donnelly. 2014. "NHS scandal chiefs launch 'reputation management' firm." in *The Telegraph*. London, UK.

> The CQC had quite a big internal agenda in terms of bringing together the predecessor organisations, creating a common culture, working out exactly how the new body was going to work internally and, ultimately—and, of course much more importantly—relate to providers in the system. *Mr Stephen Dorrell, Chair.*[947]

The expectations about the new regulator may have shaped the audit process. The NAO study, for example, 'examines how the Commission has used its resources in carrying out its quality and safety assurance work.'[948] As Greene acknowledges:

> Politics can be keenest when evaluation is intended to inform decisions about the continuation or elimination of a controversial policy[949]

There was interest in valuing the decision of merging previous regulators into one organisation. Consequently, the focus on its effectiveness in the provision of services was a central variable under the eyes of the NAO. In terms of politics, it is interesting how the different reactions about the performance of the CQC converged into a common purpose that was judging whether the organisation was fit for purpose. In a sense, the controversy around the CQC generated some pressure on audit bodies to act and to question the officials responsible of the organisation. Even the government promoted the performance of independent inquiries about the state of healthcare in England.[950] In the Select Committees inquiries, the focus of the analysis was also the capacity of the CQC to deliver. These made evident that the CQC was underperforming:

> The Committee concluded that the bias in the work of the CQC away from its core function of inspection and towards the essentially administrative task of registration represented a significant distortion of priorities. Health Committee.[951]

---

[947] Health Committee. 2011c. "Annual accountability hearing with the Care Quality Commission." London, UK. Evidence 1.

[948] National Audit Office. 2011a. "The Care Quality Commission: Regulating the quality and safety of health and adult social care." London, UK. P. 5.

[949] Greene, J. C. 2003. "War and peace ... and evaluation." *Studies in Educational Policy and Educational Philosophy* 2. P. 2.

[950] A well-known example was the Francis Report which aimed 'to examine the operation of the commissioning, supervisory and regulatory organisations and other agencies, including the culture and systems of those organisations in relation to their monitoring role at Mid Staffordshire NHS Foundation

Trust between January 2005 and March 2009 and to examine why problems at the Trust were not identified sooner, and appropriate action taken.' Robert Francis QC. 2013. "Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry." London, UK. P. 10.

[951] Health Committee. 2011c. "Annual accountability hearing with the Care Quality Commission."

> The Commission [...] has failed to fulfil this role effectively [...] As a result;
> the public are unclear what the Commission's role is and lack confidence
> that it is an effective regulator. PAC [952]

The assessment guided predominantly by values like efficiency and effectiveness showed that the tasks given to the CQC have not been properly accomplished. As an NHS official (UK-13) highlights, the focus was on making the CQC accountable for the money. As a newly created body, it was given a wider range of tasks under more restrictive conditions, i.e. less resources, new legislation.

Although the CQC was set up under the Labour government in 2009, most of its operation occurred under the Coalition Government. This echoes the ideas about 'the blame game'; the failure of the CQC can represent a scenario in which both political parties can be blamed—the Labour Party for the decision of merging three regulators (design), while the Coalition Government can be blamed for the shortcomings of the organisation (implementation). The establishment of the purpose of evaluation in the case of the CQC is controlled by key stakeholders (DoH, NAO, PAC). As a result and, according to the theory, the control exercised by these actors at this stage can have important influence on subsequent decisions about evaluation. Under the VFM scope, the work of the NAO and of Select Committees dismissed the discussion of other variables relevant in the health and social-care sector. For example, the quality of care, access to treatment, patient safety, among others helpful for analysing the role of a regulator, which represents a highly complex issue.[953]

In terms of *resources*, the NAO generates and provides the evidence for the follow-up inquiries made by the PAC. Evidence-collection is a process subject to the discretion of the NAO or Select Committees. Even though there are some standardised guidelines for submitting evidence to a Select Committee inquiry, there are discretionary spaces for staff and MPs to choose the evidence that is considered more suitable for the purposes of the study. A former member of the

CQC's Board (UK-10) acknowledged that there was a constant interaction between them and advisory groups to identify areas of improvement. Although the institutional framework allows participation in the scrutiny of public bodies, this might be constrained by different variables. Knowledge about the procedures, networking and geographical location can limit the access to some actors that might provide relevant evidence about a policy. For example, for the NAO study most of the stakeholders consulted (circa 77%) were London-based organisations, i.e. Age UK and NHS Confederation, which had better conditions for accessing this process.[954]

Access and participation are also related to the power and interests of stakeholders. For example, in the latest Health Committee's inquiry, nearly 75% of the witnesses were large organisations from the governmental sphere (DoH, CQC, Foundation Trust Network) and from the social sector, i.e. General Medical Council and The Royal College of Radiologists.[955] Some individuals also gave evidence regarding their personal experiences with the CQC's work and their concerns had already arisen in other forums.[956] This suggests that the institutional channels for accessing the process allow the participation of a range of stakeholders; most participants are organisations whose interests might be directly affected by work of the CQC. For example, the *Relatives and Residents Association*, a recurrent evidence provider to the Health Committee, expressed their concern about 'the inadequacy of the CQC's inspection methodology'[957] which is directly related to their tasks (providing information to support the selection of a care home).[958] If the criteria for evaluating a care home differ from the one used by the CQC it might generate controversy among

---

[954] National Audit Office. 2011e. "The Care Quality Commission: Regulating the quality and safety of health and adult social care. Detailed Methodology.". London, UK: National Audit Office.
[955] Health Committee. 2014a. "2013 Accountability Hearing with the Care Quality Commission." London, UK.
[956] For example, Margaret and Janet Brooks gave evidence in relation to the case of their mother who received poor medical care in a NHS Trust Hospital. They have already made their concerns public to the Nursery and Midwifery Council (NMC) and to the Parliamentary Health Service Ombudsman.
[957] Health Committee. 2013b. "Written evidence from the Relatives & Residents Association ". London, UK.
[958] This association 'is a national charity which exists for the benefit of older people in residential care, as well as their families and friends left behind at home.' The Relatives & Residents Association. 2012. "Who we are."

organisations involved in this area.

Data provided by the evaluated body is one of the most important inputs for an evaluation. As a DoH official states (UK-9), it is essential to respond to Parliament on behalf of the government about the CQC's performance. The discretion to disclosure data is limited—not annulled—by the statutory power of the NAO. Moreover, it is important to note that the CQC is not the only governmental source of information; both the DoH and the NHS produced information for the inquiries. In terms of the flow of information, the CQC appears to be a cooperative body. For example, a former Board member (UK-10) stated that the CQC staff was interviewed for the NAO study. This can have political implications too, even though finding some resistance to evaluation is not uncommon. Exhibiting the flaws of the organisation can generate political consequences.[959] For instance, Kay Sheldon, one of CQC's Board members, gave evidence to the Mid-Staffordshire Inquiry (Francis Report) about the CQC's deficiencies:

> It was a very hard decision to give evidence. I'd tried to raise the issues internally and with the Department of Health. But all I got was either denial or a sense that I was the problem. I was asked not to attend any further board meetings. I declined that request - but the meetings were very difficult.[960]

This generated different reactions within the CQC, which increased the controversy around this regulator, particularly because this member directly accused the DoH of dismissing her concerns:

> I was disturbed to hear Una O'Brien say that my concerns were "well known" to the Department and that I had contributed them to the Capability Review. This is not the case. I have stated several times to Una over the last two months (I have emails as evidence) that I have not had adequate opportunity to describe, discuss and evidence the serious concerns I have about the leadership, management and culture of CQC.[961]

This shows the extent to which politics can interfere in the performance of an evaluation or an audit. The CQC case allows looking at the discrepancies

---

[959] Geva-May, I., and W. Thorngate. 2003. "Reducing Anxiety and Resistance in Policy and Programme Evaluations. A Socio-Psychological Analysis." *Evaluation* 9(2):205-27.
[960] Dreaper, J. 2013. "CQC whistle-blower: 'I am tainted because I spoke up'." BBC News.
[961] Committee of Public Accounts. 2012. "The Care Quality Commission: Regulating the quality and safety of health and adult social care." London, UK. Evidence 37.

between stakeholders, the confrontation that arises from the process and the power that can be exercised in the pursuit of interests.

Funding is another variable to discuss in relation to resources. Regarding this, during the last six months the CQC reported having spending £911,009.26 in research services.[962] An interesting example is the study developed by the *Manchester Business School* and *The King's Fund* about the regulatory model. According to evidence submitted to the Health Committee in 2012, the cost of the study would be £50,000.[963] Between September 2013 and March 2014 the CQC paid £156,717.80 combined to both institutions under the concept of 'professional services'. However, as transparency information does not include the specific products that those services comprised it is not possible to know if these payments belong to the mentioned study. This matters because opacity about the costs of evaluation increases uncertainty in the market as both suppliers and demanders have incomplete data to make an informed decision. As it has been observed by House in the case of the United States, the dynamic that is created between commissioners and evaluators is different from any other market:

> Instead of an autonomous, impersonal market of many buyers and sellers, this particular sector of the evaluation market consists of a set of bilateral associations between the government and firms, based on specific, personal, interdependent interactions over a long period of time, what some might call an "imperfect" market from a neoclassical view.[964]

These 'imperfections' are also relevant when discussing independence and the credibility of evaluation. In relation to human resources available for evaluation and auditing, the CQC case shows the existence of a very active community of NGOs involved in the debate about the effectiveness of the CQC, and that can be very critical. After the publication of one of CQC's reports,[965] for example, the

---

[962] This is the total amount of the spending reported by the CQC under the label of 'research services' from September 2014 to February 2015. Source: Care Quality Commission. 2015b. "Spending transparency." London, UK.

[963] Health Committee. 2012. "2012 Accountability Hearing with the Care Quality Commission." London, UK. Evidence 49.

[964] House, E. R. 1997. "Evaluation in the Government Marketplace." *American Journal of Evaluation* 18(1):37-48. P. 39.

[965] Care Quality Commission. 2011b. "Dignity and Nutrition Inspection Programme. National Overview.". London, UK.

Chief Executive of *The Patients Association* stated:

> Yet another report to add to the plethora of other reports, all highlighting the same issues around poor care of elderly patients in hospital. How many reports do we have to look at and how many times the public need to hear about this before the right action does is taken. What will it take to bring about the change that is needed so that elderly people get treated in Hospitals with the respect and dignity they deserve.[966]

This was seconded by the Age UK's Chief Executive:

> Age UK renews its call to Government to take the responsible course by appropriately funding and guaranteeing fair access to high quality health and social care services [...] this report[967] shows the dire consequences of a social care system that has been under increasing financial pressure over the last eight years and in many areas is now financially stripped to the bone.[968]

The space for NGOs to discuss public issues is linked to their level of independence and their distance to audit bodies, as a member of a social-care charity (UK-6) highlights: 'We can say what we want. If you're getting funding from government you can't possibly criticise government.' There is availability of information, financial resources, and experts to assess the CQC's performance. In terms of the hidden politics of evaluation, the relevance of resources for evaluation and auditing is how these are allocated and how this impacts the level of power and control of stakeholders.

The *process* of evaluation appears to be the most complex area to analyse, due to the multiplicity of factors that intervene. One of the most political issues associated with this element is the interaction between stakeholders. For example, the configuration of the PAC establishes the basis for making the government accountable and it is a major component of the institutional framework that determines how each stakeholder can participate. Without the obligations and prerogatives given to the NAO and the political support provided by the PAC, the establishment of an accountability system would not be possible or effective. The fact that the PAC's Chairman is a member of the leading opposition limits the potential influence of the government in its

---

[966] The Patients Association. 2011. "Patients Association reacts to results of CQC." London, UK.
[967] Care Quality Commission. 2012b. "The state of health care and adult social care in England ". London, UK.
[968] Age UK. 2012. "Response to CQC 'State of Care' Report." London, UK.

work,[969] a logic that also underpins the appointment of the C&AG.[970]

> He's appointed through a process that involves the Prime Minister of the day and the Chairman of the PAC of the day, who is always a member of the leading opposition party, so he's appointed, so he can only be appointed by consensus of both sides, but then his independence is guaranteed by statute, the government can't influence what he writes in these reports (UK-12).

The relationship between the NAO and the PAC is also framed by the institutional framework. This gives this audit body independence 'to pursue an audit agenda and report publicly without external influence, even from MPs.'[971] Despite the distance between both stakeholders formally established, it is not possible to ignore the fact that politics is a very important component of their work. The work of both bodies can generate important political reactions, as multiple interests can be affected by their inquiries. The institutional framework provides some warranties in relation to the Chair of the PAC. This aims to establish a counterpart to government for the discussion about the implementation of policy and how resources are spent. It represents an important characteristic of the UK case, considering that evaluation and audit outputs can represent a threat for some stakeholders.[972] Therefore, it is important to acknowledge the politics of these activities in relation to the interests pursued by each stakeholder.[973]

The openness to the involvement of external stakeholders is another variable to discuss. This is left to the discretion of the actor that controls the process. In this case, it appears that the studies about the CQC involved the integration of

---

[969] For example, the last Chair of the PAC was Margaret Hodge, a member of the Labour Party.

[970] According to the Public Accounts Commission, 'the C&AG's independence is preserved even from the House by the fact that he or she has, by statute, complete discretion in the discharge of his functions.' The Public Accounts Commission. 2008. "Corporate governance of the National Audit Office: response to John Tiner's Review." London, UK: House of Commons.

[971] Dunleavy, P., C. Gilson, S. Bastow, and J. Tinkler. 2009. "The National Audit Office, the Public Accounts Committee and the risk landscape in UK public policy." London, UK: The Risk and Regulation Advisory Council. P. 4.

[972] Datta, L. E. 2011. "Politics and Evaluation: More Than Methodology." *American Journal of Evaluation* 32(2):273-94, Hatry, H. P., J. S. Wholey, and K. E. Newcomer. 2010. "Evaluation Challenges, Issues and Trends." in *Handbook of Practical Program Evaluation*, edited by Harry P. Hatry, Joseph S. Wholey, and Kathryn E. Newcomer. US: Jossey-Bass.

[973] Lipsky, D. B., R. L. Seeber, A. C. Avgar, and R. M. Scanza. 2007. "Managing the Politics of Evaluation: Lessons from the Evaluation of ADR Programs." Pp. 116-29 in *Proceedings of the Fift-Ninth Annual Meeting of the Labor and Employment Relations Association*.

different views about its performance. This is encouraged by the institutional framework and has given the PAC an important place in the accountability process and more leverage to its recommendations. A CQC official states (UK-1):

> I think they've all over the past 3-4 years have really developed a reputation for holding the government to account for the ways it is delivering its policies that would include all the people that I need to account and I see that as a good thing.

The PAC is an important element of the audit system in the UK, particularly because it gives political support to the work of the C&AG. Moreover, the political power of the Chair also contributes to the credibility and notoriety of the PAC's inquiry and it has modified the traditional relationship with other actors, according to a former clerk of the PAC (UK-7):

> Until 2010 the work of the PAC was purely the dialogue between the committee and the department, so you'd have oral evidence session only with the officials from the department and there would not be politicians, there would be permanent officials, the permanent secretary and other staff, in the last four years under the chair of Margaret Hodge they have included evidence from many more outside sources.

The inclusion of new actors in the discussion about the performance of public organisations shows the way in which the institutional framework can allow or restrict this access. A relevant point here is the leadership of the former Chair, Margaret Hodge MP, and how she gave the PAC important notoriety. On the one hand, the former Chair was criticised for being severe and harsh in her questionings—even to the point of being accused of bullying and abuse.[974] As a Labour MP declared: 'When she's got her teeth into something, she won't let go of the bone.'[975] On the other hand, the political strength of the PAC also produced confrontation with other Select Committees, for example, in the case of the tax avoidance investigation[976] about which …

> … Members of the Treasury Committee clearly regarded the tax avoidance investigation as an armed incursion into their territory,

---

[974] Syal, R., and M. Sweney. 2015. "Margaret Hodge accused of 'abusive and bullying' attack on BBC Trust chair." in *The Guardian*. London, UK.

[975] Mason, R. Ibid."Brief, brutal and very public: there's more to Margaret Hodge's grillings than dramatics."

[976] Committee of Public Accounts. 2013b. "Tax avoidance: The role of large accountancy firms ". London, UK: House of Commons.

and complain that they would have made a better job of it.[977]

This underlines the role of the PAC in the UK's landscape. The political support of the House of Commons not only makes explicit the distance between Select Committees and the government, but also legitimises the work of the NAO as the external body that scrutinises governmental accounts. This becomes crucial when addressing issues like independence and credibility.[978] Currently, under a new Chair, new questions arise about the way in which PAC will hold the government accountable.[979]

The coexistence of different bodies scrutinising the performance of the CQC has generated a greater pressure for this regulator to overcome its initial failures. As a CQC official states (UK-1): 'The issue is being clear about how we are discharging those multiple accountabilities.' The existence of a plethora of bodies that make this regulator accountable shows that the UK's institutional framework is characterised by the interaction of multiple actors whose action might be guided by a different rationale. In the opinion of a former member of the Board (UK-10) there is an effective checks and balances system that also involves the participation of external stakeholders who can be affected by the work of this regulator.

The CQC case shows that the rigorous scrutiny has been a constant, even though responding to it is 'not comfortable[…] but actually it is necessary' (UK-1). The relevance of the existence of these bodies is that 'it is important that the public know that money is being used properly' (UK-10). The openness to scrutiny and evaluation is an important variable that can facilitate or difficult the process. Regarding this, a Health Committee staff member (UK-2) states:

---

[977] D'Arcy, M. 2015. "Cracks in the PAC?". London, UK: BBC News.
[978] Brown, A. B., and J. A. Klerman. 2012. "Independent Evaluation: Insights from Public Accounting." *Evaluation Review* 36(3):186-219.
[979] The election of Meg Hillier MP as the new Chair of the PAC has been criticised as she is not considered a strong figure like the former Chair Margaret Hodge and in relation to other Labour MPs that were participating in the election (Helen Goodman, Gisela Stuart). Regarding the outcome of the election, a journalist states "they could have chosen Gisela Stuart, who is known for her contempt for 'rudderless leaders who drift with the political wind', or Helen Goodman, who would have brought specialist knowledge to the post, as a former Treasury civil servant. Instead, they chose Meg Hillier, whose greatest distinction is that Ed Miliband sacked her from the shadow Cabinet because she was hopeless." McSmith, A. 2015. "Andy McSmith's Diary: So that's how Meg Hillier became Margaret Hodge's successor." in *The Independent*. London, UK.

> They are used to the process now and they can't avoid it and if you are not helpful what it will do is to create impression on most members of the Committee that they are trying to avoid scrutiny and if you create the impression in front of the Committee then you are hiding things you are going to make them look harder and make them more sceptical about what they do find.

In terms of politics, this resonates with the concept of evaluation culture (see Chapter 2). According to Højlund, organisations interested in increasing efficiency are more likely to have a culture of evaluation, because its legitimisation is based on outputs.[980] In this scenario, criticisms to the CQC might influence how this organisation embraces evaluation as a part of its daily work. Moreover, the institutional framework enforces the CQC to heed the recommendations of audit bodies. Audits and evaluations are important political tools to promote change both internally and externally.[981] For instance, the Health Committee inquiry fostered the development of a study 'to show how evidence and research can be used by CQC to evaluate how well its current regulatory arrangements in health and social care work.'[982]

Despite the interaction between evaluators and evaluated bodies being constrained by the institutional framework, this does not completely eliminate discretionary spaces. An example of this (also observed in the Cancer Strategy case) is the 'clearance stage' that occurs during a VFM study. Here, the audited body can comment on the NAO's report and even though it was emphasised by interviewees that evaluated bodies cannot make changes to the content, one of them recognises:

> You get to a sort of nuances about possibly tone so sometimes you may argue over wording, so you're still saying the same thing but how you say it can change, so those can be more difficult conversations, because it's less about facts and more about tone. (UK-3).

A former PAC's clerk (UK-7) agrees, 'There's often a process of discussion in the Committee about whether the wording of the report is right and this

---

[980] Højlund, S. 2014. "Evaluation use in the organizational context – changing focus to improve theory." *Evaluation* 20(1):26-43.

[981] Rist, R. C., M. H. Boily, and F. Martin. 2011. *Influencing Change. Building Evaluation Capacity to Strengthen Governance*. Washington, D.C., US: The International Bank for Reconstruction and Development/The World Bank.

[982] Walshe, K., and D. Phipps. 2013. "Developing a strategic framework to guide the Care Quality Commission's programme of evaluation." London, UK: The University of Manchester.

occasionally takes a party political tension.' These statements evince informal mechanisms that public bodies can use to influence reports. The degree to which this is possible is unclear, but there are documented cases in which a report has been delayed because of disagreements at this stage.[983] While there is no explicit intention to manipulate the findings of a report, changes in the tone and wording can become mechanisms to make a criticism softer, to be less straightforward about a particular issue, and to generate less political reactions. To rephrase a statement by turning a 'weakness' into an 'opportunity for improvement' can involve making a judgement more subtle.[984] This has been a recurrent issue in the discussions about the political nature of evaluation. Petitions for changing the wording of findings are constantly found in the interaction between evaluators and evaluands.[985]

The inclusion of non-governmental actors in the assessment is also part of the process. As in the previous case, there are institutional mechanisms for the incorporation of different views, both in internal evaluation activities as well as in inquiries led by the NAO or Select Committees. Regarding this, a researcher of a social-care charity (UK-6) acknowledges:

> If we feel that there's a theme that is coming across again and again that it affects majority of people, then we would influence the policy and we would influence the politicians.

To exercise their influence, organisations use official information as well as data produced by them.[986] Here, the role of evaluators becomes more relevant. They have the capacity to promote the inclusion of different views, yet this is constrained by the power of the actor who commissions the evaluation or even

---

[983] This was the case of the NAO's report on the NHS's National Programme for IT. See: Ritter, T. 2007. "National Audit Office - the murky "clearance" process."

[984] This has also been observed and documented previously by different authors. See: Dunleavy, P., C. Gilson, S. Bastow, and J. Tinkler. 2009. "The National Audit Office, the Public Accounts Committee and the risk landscape in UK public policy." London, UK: The Risk and Regulation Advisory Council, Sharma, N. 2007. "Interactions and Interrogations: Negotiating and Performing Value for Money Reports." *Financial Accountability & Management* 23(3):289-311, White, F. 1999. *Audit, Accountability and Government.* Oxford, UK: Clarendon Press.

[985] The LSE GV314 Group. 2014. "Evaluation under contract: Government pressure and the production of policy research." *Public Administration* 92(1):224-39.

[986] The Patients Association, for example, has a Helpline for patients and their relatives in which they provide information about different aspects of the health system as well as advice about any concerns that patients might have in relation to healthcare. The data derived from this helpline informs several of the projects carried out by this organisation.

by evaluators, as one of them (UK-14) describes:

> We would seek the views of the people we think it would be helpful to speak to, of course […] we would always have to check with the client, this is what we propose to do and this it who we think it would be useful to speak to, of course the client would have an idea because they have inside knowledge on what's going on, the politics, they might have a view, however if they don't want you to speak to somebody in a way that's data, it says something.

For example, according to the methodology of the study undertaken by the *Manchester Business School* and *The King's Fund*, the CQC participated in the process at different moments, 'including commenting on the selection of inspections to observe and on the design of interview and survey questions prior to finalisation.'[987] This is again related to the proximity between the evaluated body and other stakeholders, because it raises concerns about independence and credibility. For example, some NGOs that have participated as witnesses in Select Committee inquiries have also been commissioned to research products, an issue that can create conflicts of interest.[988] If an organisation plays a double role, both as witness and as provider of services for the audited body, this can have negative effects on the credibility of the evaluation. What the process of evaluation and auditing in the CQC case suggests is that when the evaluated body is embedded in a highly politicised context, this will also affect the interaction between stakeholders and, consequently, the outputs and outcomes of evaluation.

The hidden politics of evaluation also concerns the *outputs.* The products derived from the evaluation and audit processes to the CQC make evident those areas that need improvement. The influence of politics can be observed here in the use and dissemination of findings. According to Stevens, evidence can be used to support an idea previously developed by a stakeholder.[989] In the CQC case, the utilisation of findings can be more clearly seen in the promotion of

---

[987] Walshe, K., R. Addicott, A. Boyd, R. Robertson, and S. Ross. 2014. "Evaluating the Care Quality Commission's acute hospital regulatory model: Final Report." P. 40.

[988] This was observed in the cases of The Patients Association, the Relatives and Residents Association and Carers UK who have given evidence in inquiries of the CQC and also appear as providers of research services in the transparency data published by the CQC.

[989] Stevens, A. 2011. "Telling Policy Stories: An Ethnographic Study of the Use of Evidence in Policy-making in the UK." *Journal of Social Policy* 40(02):237-55.

organisational changes. The way in which evaluation and audit outputs are integrated in the logics of organisations supports the argument made by Hansson about how these mechanisms are a consequence of the logic imposed by the NPM:

> The growing interaction of evaluation in the processes of organization and management is a forceful demonstration of how evaluation is becoming an integrated part of the organizational environment under the new public management system.[990]

Indeed, this new dynamic has forced organisations to embrace mechanisms of control to generate evidence about its performance and to have elements that justify its existence. An important element derived from this is the need to be politically profitable. For the government in office this serves the purpose of proving the effectiveness of its policies, while for the opposition it generates the opportunity to make evident the need for a change. In both cases, positive and negative findings derived from evaluations or audits can be used by different stakeholders. One of the most evident examples in this case is the changes that took place within the organisation's governance. The NAO's report had already pointed out that the Board had not been able 'to define success.'[991] Consequently, this made evident a lack of direction that became more notorious at the PAC's hearing, for which one MP highlighted:

> One of its own board members has gone on record about the repeated failures of delivery, governance and effective leadership.[992]

The deficiencies in the CQC's governance were also publicly disseminated by the media.[993] This might have produced the contextual factors necessary to promote use. As Green et al. discuss, context can represent an important factor that facilitates the use of knowledge;[994] in this case, the evidence generated by

---

[990] Hansson, F. 2006. "Organizational Use of Evaluations: Governance and Control in Research Evaluation." *Evaluation* 12(2):159-78. P. 163.

[991] National Audit Office. 2011a. "The Care Quality Commission: Regulating the quality and safety of health and adult social care." London, UK. P. 15.

[992] Committee of Public Accounts. 2012. "The Care Quality Commission: Regulating the quality and safety of health and adult social care." London, UK. Evidence 24.

[993] See for example: Campbell, D. 2011. "Care Quality Commission's leaders 'lack the necessary skills' " in *The Guardian*. London, UK, Swinford, S., and C. Gammell. 2011. "Warnings of care home 'were ignored for months'." in *The Telegraph*. London, UK.

[994] Green, L. W., J. M. Ottoson, C. Garcia, and R. A. Hiatt. 2009. "Diffusion theory and knowledge dissemination, utilization, and integration in public health." *Annu Rev Public Health* 30:151-74.

evaluation and audit processes. The CQC offers an interesting case study in terms of how politics can shape not only the evaluation and audit agenda but also the use of findings. It is not possible to determine exactly the linkage between the political notoriety of the CQC and how evaluation and auditing fostered changes in the organisation, but it is clear that the attention given to this regulator was driver for the transformation of the CQC.

In organisational terms, for example, the intense scrutiny resulted later in the resignation of the former Chair and Chief Executive,[995] and it led to the appointment of new Board members.[996] An interesting point to highlight is the inclusion of Robert Francis QC, who led the *Francis Report*, and also the fact that Kay Sheldon, who denounced failings in the governance of the CQC, was the only member who remained in the Board after these changes took place.[997]

Another important issue to discuss is the capacity of bodies like the NAO or Select Committees to enforce public bodies to heed recommendations. The evidence suggests that there is no formal obligation for public bodies to follow recommendations made by bodies such as the NAO or Select Committees. However, they are expected to respond to those recommendations and to argue in case of dismissal. The CQC acknowledges:

> The external scrutiny raised questions about our performance and role as the health and social care regulator. The Board acknowledges many of the criticisms that have been made about the effectiveness of the organisation during its first three years of its life […] over the next year we will reset the strategic direction of the organisation, measure CQC's impact and effectiveness, and hold the Executive to account for delivery.[998]

This leads to making two major points. The first one is that the institutional framework allows public bodies to either agree or not with recommendations. For example, in relation to the registration problems faced by the CQC highlighted by the PAC, 'The Department does not agree with the Committee's

---

[995] See: BBC News. 2012. "Dame Jo Williams: Care regulator chair to resign." London, UK. and Watt, H. 2012. "Head of Care Quality Commission quits following damning report." in *The Telegraph*. London, UK.

[996] Care Quality Commission. 2014b. "CQC welcomes new Board members." London, UK.

[997] See: —. 2015a. "Board Members." London, UK.

[998] —. 2012a. "Performance and Capability Review. CQC Action Plan." London, UK. P. 2.

conclusion that it has only recently begun to take action.'[999] This is an important element to promote the public debate and to engage stakeholders in the process.[1000] It suggests that there is room for discussion and not a unilateral imposition of judgement. The second is that the institutional framework also generates ambiguity about how recommendations will be implemented. The discretionary space given to public bodies can be observed in the following statement made by a CQC official (UK-1):

> One is: Do you accept the substance of it, they are not making a recommendation about how they think that recommendation should be taken forward. I think we have to take a view about not just the what we've been told to do, but how, so we might accept the 'what' but think in different 'hows' to deliver a recommendation [...] I mean you accept the spirit of the recommendation—not the detail of it.

The way in which recommendations are followed show how political power is exercised. For example, derived from the PAC's finding: 'There are inconsistencies in the judgements of individual inspectors and in the Commission's approach to enforcement.'[1001] The DoH and the CQC carried out important modifications to the regulatory model that included the performance of an evaluation study[1002] and a public consultation with stakeholders.[1003] This suggests that the focus of the CQC was on the development of a new regulatory model considering that this affects the interests of different stakeholders, i.e. providers, GPs, and the adoption of changes in this area had to be driven more carefully. For example, the increase of regulatory fees[1004] and the public display of CQC ratings[1005] are particularly sharp issues for which the reactions of key stakeholders can already be seen:

---

[999] HM Treasury. 2012b. "Treasuy Minutes. Government responses on the Fifty Sixth, the Seventy First, the Seventy Sixth and the Seventy Eighth Reports from the Committee of Public Accounts: Session 2010-12.". London, UK. P. 16.

[1000] Adams, A. E., N. A. Nnawulezi, and L. Vandenberg. 2014. ""Expectations to Change" (E2C): A Participatory Method for Facilitating Stakeholder Engagement With Evaluation Findings." *American Journal of Evaluation*.

[1001] Committee of Public Accounts. 2012. "The Care Quality Commission: Regulating the quality and safety of health and adult social care." London, UK. P. 6.

[1002] Walshe, K., R. Addicott, A. Boyd, R. Robertson, and S. Ross. 2014. "Evaluating the Care Quality Commission's acute hospital regulatory model: Final Report."

[1003] Care Quality Commission. 2013g. "A new start. Consultation on the changes to the way the CQC regulates, inspects and monitors care." London, UK.

[1004] Care Home. 2015. "CQC announce nine per cent increase in regulatory fees." London, UK.

[1005] Warrington Guardian. 2015. "CQC ratings to be displayed publically at health and care providers."

> It is extremely disappointing that the views of 80% of the respondents to the consultation have been ignored and GP practices will be facing yet another increase in CQC fees.–*Deputy Chairman, General Practitioners Committee.*[1006]

In terms of the implementation of recommendations, the relationship between the CQC and the DoH plays an important role. It is clear that the DoH sets the policy and the CQC implements it, so there is a strong bond between both organisations that might influence the way in which recommendations are instrumented. A DoH official (UK-9) explains that their role is 'to share the objective', and also 'to make explicit that it is a separate body with independent judgement and against the perception of political interference'; however, this seems to produce a conflict of interest for the DoH to respond to Parliament on behalf of the CQC, while simultaneously it aims to promote the independence of this body. In the context of the fieldwork undertaken for this research, the intention of not being 'too open' was perceived in the discussion of this issue, in the sense that this might put interviewees in an awkward position. This reveals the political interests at stake involved in the duties of the CQC. Despite the incorporation of different mechanisms for the obtaining of evidence to support decision-making, the political factor remains inherent to all stages of policymaking, including evaluation.

The other element to discuss is the dissemination of findings. The institutional framework for auditing, for instance, promotes the transparency the NAO's work and of the Select Committee hearings. This ensures that documents derived from these processes are made public. In relation to politics, this represents an important opportunity for stakeholders to exercise their power. As Gaarder and Briceño observe:

> If the entity overseeing the public evaluation efforts is vulnerable to political changes, e.g. if its existence is politically determined, its budget allocation, or its staff, then the independence in mind and appearance principles may be compromised, as it may be or feel under pressure to report successes only.[1007]

This case is useful for identifying how the institutional framework can limit the

---

[1006] Millett, D. 2015. "CQC's 9% fee hike is a 'financial blow' to GP practices." London, UK.

[1007] Gaarder, M. M., and B. Briceño. 2010. "Institutionalisation of Government Evaluation: Balancing Trade-Offs." International Initiative for Impact Evaluation (3ie). P. 6.

political influence in evaluation and audit processes. The independence conferred to the NAO, in addition to the support given to this body by Parliament, prevents that its work might be negatively influenced by other stakeholders. Moreover, the case of the CQC also illustrates the benefits of having an active community of external stakeholders interested in the discussion of the findings. The reports derived from Select Committees' inquiries, for example, have been discussed by external stakeholders.[1008] This helps reducing the discretion in the selective dissemination of findings, as there are more stakeholders involved in the follow-up of an evaluation. The extent to which stakeholders participate in this process can range from disseminating the results of a study to actively engage in a deeper discussion of the results.[1009]

Finally, the *outcomes* of evaluation are difficult to determine because of the multiple factors that might have influenced the changes produced in the long term and that 'cannot be easily divorced from procedural and normative measures.'[1010] This leads to questioning: *Where does the line stand between what can be considered the effects of evaluation and contextual factors such as political momentum and economic drivers?* Furthermore, the case of the CQC is particularly interesting because of its evolution across time and the events that have affected its performance. First, the origin of the CQC as the merger of other regulators has evidently impacted the early stages of its operation. Unrealistic performance targets set for this regulator, as well as a poor organisational capacity, clearly undermined its capacity to successfully fulfil its task and to level up to 'the expectations of the public.'[1011]

The public attention brought to the CQC by events like Winterbourne View increased the pressure on this organisation and led to a more severe scrutiny.

[1008] The Relatives & Residents Association. 2013. "Care Quality Commission not yet successfully defined its core purpose or earned public confidence ". London, UK.

[1009] For example, recently The King's Fund disseminated the report published by the NAO about the improvement of cancer services and outcomes in England, while McMillan Cancer Support released a formal response to this report. See: Health Management and Policy Alert. 2015. "Progress In Improving Cancer Services And Outcomes In England." London, UK. McMillan Cancer Support. 2015. "Macmillan Cancer Support responds to NAO report on cancer services." London, UK.

[1010] Fukuyama, F. 2013. "What Is Governance?" *Governance* 26(3):347-68. P. 356.

[1011] Gary Fitzgerald, Chief Executive, Action on Elder Abuse. Committee of Public Accounts. 2012. "The Care Quality Commission: Regulating the quality and safety of health and adult social care." London, UK. Evidence 4.

However, changes in the organisation and its policies cannot be automatically associated with these processes. Evaluating outcomes is challenging because there is no 'logic model that helps program managers see the relationship among inputs, processes, and outputs and the key roles played by formative feedback and contextual variables in managing for results.'[1012] In this particular case:

> I think the next difficulty then is what is caused by what we do and what is a correlation, improvements take place in complicated context organisations for lots of different reasons, so actually saying it's all caused by what we do is probably too simplistic, so actually we need some quite rich and sophisticated research and evaluation tools to work out what it is that we do that is unique in driving improvements, is distinct from what it is others in the system do that will drive improvements (UK-1).

The effect of the assessment in the broader context of policy cannot be isolated, and yet some nuances of its influence can be identified, as a member of the Health Committee staff explains:

> The accountability hearings worked pretty successful overall in identifying what the CQC was doing wrong and what they needed to do right and politically as well it provided, the act that this Committee was doing these reports, provided the momentum for the government to go away and change what the CQC was doing and to change the regulation, the regulatory work that they do, and ultimately is a good thing because they are a better organisation now than they were three years ago... (UK-2).

The case of the CQC provides a good example of how evaluation can become a powerful tool to promote change in an organisation. The modifications in the CQC's governance appear to have been directed to build more credibility within the public landscape. The appointment of a new Chief Executive in 2012 and a new Chair in 2013 generated *momentum* for its reorganisation and the establishment of new priorities just before the elections, which might bring new changes to the CQC, i.e. even a potential merging with the Human Fertilisation and Embryology Authority. On the eve of the elections, a CQC official (UK-1) reflects on the implications of these changes:

> At the minute we're just coming to the end of the first year of a three-year reform journey and I think they should let us complete the three years of the reform journey before we start looking at additional things.

[1012] Schalock, R. L., and G. S. Bonham. 2003. "Measuring outcomes and managing for results." *Evaluation and Program Planning* 26(3):229-35. P. 230.

Furthermore, the appointment of new Board members can be seen as an attempt to rebuild public trust. As stated, Robert Francis QC was recently appointed as a non-executive Board member. He is well known for being critical about the role of this body and also an advocate for patients' rights.[1013] The inclusion of such a figure can bring to the CQC more credibility and it sends out the message that the failings that occurred in the past will not be repeated. The CQC case makes evident that the appraisal of an organisation's performance can be driven by multiple purposes. The most relevant is the need for legitimisation of the decision about merging the regulation of health and social care in a unique body. The initial difficulties faced within the CQC and the criticism about its response to highly publicised incidents did play a role in the way the organisation was assessed. Its relationship with the DoH and the way its credibility was affected also influenced the process. The CQC offers a case study in which assessment is underpinned by subjective elements affected by politics. In order to establish clear elements for comparison, the next sections are dedicated to the analysis of the Mexican case studies.

[1013] Robert Francis QC was Chair of the Mid Staffordshire NHS Foundation Trust Public Inquiry and he published a wide-spread report about the deficiencies found in this trust in patient care. He is also Honorary President of The Patients Association, an NGO that advocates for patients' rights and that has a close relationship with the CQC as a provider of evidence. According to a member of this association (UK-6), 'we have an arrangement with CQC where they've asked us to give them broad top-level findings from our helpline, I think we do it every week, so we send them information saying, you know, somebody was trying to do something and they weren't able to get answer from the trust…' The incorporation of Robert Francis to the CQC might affect this relationship in terms of independence.

6.2 Main Findings of the Mexican Case

The characteristics of the institutional framework in which evaluation and audit take place in Mexico are also useful for analysing the role of politics in these activities. This section discusses the most relevant findings of the Mexican case through the thematic framework developed in Chapter 2.

The description of the rules and procedures that guide evaluation in Mexico (see Chapter 5) represents a good point of departure for looking at how these rules affect its design and implementation. The Federal Government has fostered the development of evaluation capabilities for the institutionalisation of this activity at the national level. So far, these efforts have developed into the systematic evaluation of a wide range of policies and programmes. The audit process undertaken by the SAO represents an important advancement for the establishment of an accountability system and for the inclusion of an external view about the effectiveness of government and the use of public funds. Evaluation and audit tools are affected by the political context both in their processes and outcomes, as it is discussed in the next section.

6.2.1 The Institutional Framework and the Hidden Politics of Evaluation in Mexico

The first element that stands out in the Mexican case is its highly centralised evaluation policy. This comprises two major processes: internal evaluation and external audits made by the SAO on behalf of the Chamber of Deputies. The allocation of power promoted by the institutional framework affects the characteristics of the evaluation and audit systems implemented and the way in which participants interact (Figure 6.5).

Figure 6.5 Main Characteristics of the Institutional Framework for Evaluation in Mexico

High level of
centralization and
excessive control of the
Executive

Limited scope of
evaluations

Weak audit
body

Interference with
independence
and credibility

Reduced market
of evaluators
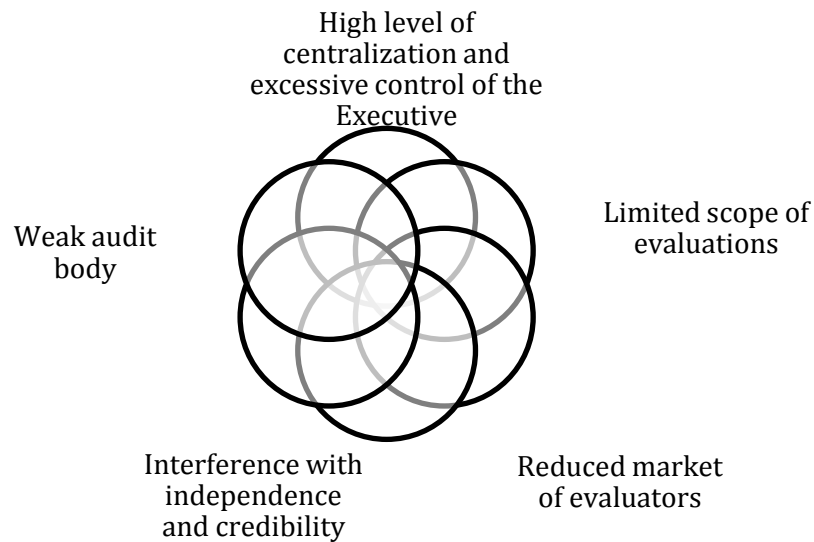
Unlike the UK case, in Mexico the Federal Government exercises strong control over the evaluation process. Most evaluations are based on standardised instruments (mostly developed by CONEVAL) that leave a very narrow space for evaluators to work outside these frameworks. In line with some of the secondary questions of this research, this element aids acknowledging the kind of evaluation instruments that have been developed in this country and which outputs can be expected from these.[1014] Design evaluation, for instance, is based on a model that not only sets the questions that the programme must answer, but also the criteria for the evaluator to assess these responses.[1015] This produces two major consequences:

- It allows obtaining standardised information about programmes that can be more easily systematised and compared.[1016]
- It simplifies the performance of evaluations through the use of models designed for this purpose. However, this limits evaluation to a reduced range of variables.

This has been a central issue in the discussion about the effectiveness of the Mexican evaluation policy, as some interviewees recognise:

[1014] This relates back to the question, *Does the evaluation policy, in both countries produce relevant and convenient information that can be used for the policymaking process?* The nature of the instruments promoted within an evaluation policy can provide insights about the kind of outputs that can be obtained.

[1015] National Council for the Evaluation of the Social Development Policy. 2015. "Terms of Reference Model for the Design Evaluation." Mexico City.

[1016] CONEVAL, for instance, developed an instrument in which it is possible to have a synthetic view of the performance of federal programmes. Source: —. 2014d. "Monitoring Record Card."

> I can see as a serious problem of these formats that a poor evaluator can stick exclusively to fit the criteria [...] but it is not an evaluation that generates added value (MX-1).

> There is a questioning about the CONEVAL's evaluation model because they use secondary information and there are all desk-based studies that do not necessarily reflect all process, for instance, even though they claim to make processes evaluation, they are evaluating lots of issues regarding design, planning [...] of course in the health sector you need to carefully select indicators to observe the direct impact on the population's health (MX-5).

This control over evaluation instruments can also be observed in the *Consistency and Results Evaluation.* This model focuses on the analysis of six themes: design, strategic planning, operation, coverage, beneficiaries' perception, and results.[1017] A programme that has been assessed with this instrument cannot provide any other information, i.e. equity or progressivity, outside this framework. This resonates with the idea that 'judging programs against pre-determined targets may give the wrong signals: the targets may be irrelevant, overambitious or exceedingly modest.'[1018] But at the same time, it makes evident that the actor who has control over the process has also the power to make those decisions according to his interests. Under this logic, an increase on the coverage could be seen as a positive finding, and yet, this would not provide enough elements to determine if a policy is working or not.

Despite the rigidity of the evaluation process and the limited access to external stakeholders, evaluation is not exempted from political interference, particularly in the case of programmes that have reached an important level of notoriety in the public landscape. A Ministry of Health official (MX-4) identifies that 'There are interests and nuances; at the end there are programmes that are protected.' This suggests that the inflexibility of the evaluation policy can be an attempt of the Federal Government to maintain evaluation at a level that does not represent a menace to their interests. This is consistent with what Behrman and Skoufias state:

> Those who institute and advocate particular programs often are convinced, a priori, that they are likely to be successful in attaining their stated

---

[1017] —. 2011b. "Terms of Reference Model for the Consistency and Results Evaluation." Mexico City.

[1018] Picciotto, R. 1999. "Towards an Economics of Evaluation." *Evaluation* 5(1):7-22. P. 16.

objectives, which is why they are strong advocates in the first place.[1019]

The political influence in evaluation can be observed at different stages of the process as it will be discussed through the case studies. Overall, the most evident influence perceivable is the one exercised by the Federal Government in the control of evaluation instruments and the incentives for maintaining a small market of evaluators.

The second element concerns the scope of the evaluation policy. This has been mostly constrained to the programmatic level and there is little debate and analysis of evaluation outputs. This was a problem detected almost since the inception of the evaluation policy. An NGO involved in public policy recognised this as one of the most evident weaknesses of the system:

> There are no institutional and normative bases for performing an integrated and systematic evaluation [...] conciliating the macro (policies), meso (programmes and organisations) and micro (public officers' performance) levels [...] there is the risk to have incomplete information for decision-making, budget allocation and accountability.[1020]

It appears that the evaluation policy operating has not been able to transcend its role as a managerial activity to become an active part of the discussion about government effectiveness and accountability. This will be explained in more detail in the following sections as it is a crucial aspect of the use of outputs.

The third element worth discussing is the market of evaluators. Its development has been the result of the rise in the demand for evaluation after the institutionalisation of this activity.[1021] At first, there was only 'a very reduced pool of evaluators' (MX-24) that began working for government ministries. This dynamic is atypical, because 'The limited supply of evaluation services means that some putative advantages of contracting out (e.g. reducing bureaucratic control, cutting costs, and speeding operations) may not obtain.'[1022] Moreover, it

---

[1019] Behrman, J. R., and E. Skoufias. 2006. "Mitigating Myths about Policy Effectiveness: Evaluation of Mexico's Antipoverty and Human Resource Investment Program." *The ANNALS of the American Academy of Political and Social Science* 606(1):244-75. P. 245.

[1020] González Arredondo, A. (Ed.). 2008. *¿Gobernar por resultados? Implicaciones de la política de evaluación del desempeño del gobierno mexicano*. Mexico City: GESOC, A.C. P. 241.

[1021] Feinstein, O., and G. Hernández. 2008. "El Papel de la evaluación en México." Banco Mundial.

[1022] House, E. R. 1997. "Evaluation in the Government Marketplace." *American Journal of Evaluation* 18(1):37-48. P. 40.

generates negative incentives in terms of independence and credibility because it favours a codependent relationship between commissioners and evaluators. As an academic (MX-10) observes:

> My perception is that because the market is reduced the independence of both commissioners and evaluators is limited [...] organisations do not have enough options to choose from [...] and even though evaluators make their job right this does not guarantee that they will be hired again so that does not necessarily mean that they might be captured or compromised to say things that are not true, but they might be more subtle in the way they say them.

Even though the institutional framework gives public organisations the prerogative to commission evaluations to external parties through public bidding, this does not eliminate the discretionary spaces in the process and its impact on the outputs of this activity.

A fourth element of discussion regards the conflicts associated with the external nature of evaluation. The model implemented by the Federal Government produces important conflicts in terms of independence and credibility. On the one hand, the control of methodological aspects of evaluation limits the work of evaluators. On the other hand, the narrow market existence creates perverse incentives to perpetuate this situation. This makes the Mexican case an interesting object of study because the institutional framework favours the 'simulation' of an external approach to evaluation. This means that despite evaluations being commissioned to external parties this does not necessarily mean that these studies are independent, because the government controls the key decisions of the process. Consequently, at a discursive level, the external nature of evaluation is seen as a safeguard for its independence, while in *praxis,* this independence is constrained by the rules and procedures imposed centrally. According to a Ministry of Health official (MX-5): 'Somehow it is the government evaluating itself and there is an important credibility issue.'

The last element to discuss revolves around the audit system. The SAO's work focuses on the identification of procedural deviations and financial reviews. Nonetheless, it has very little impact in terms of accountability and

policymaking.[1023] One of the most recurrent criticisms of the SAO is precisely its lack of 'teeth' to enforce the follow-up of recommendations and to sanction: 'Without power for establishing direct sanctions, criminal complaints made by the SAO to the General Attorney of the Republic are diluted in time and do not proceed.'[1024] Despite its relationship with the Legislative, the SAO does not appear to be appropriately supported by the Chamber of Deputies. As a member of the SAO (MX-12) states: 'We believe performance auditing is a valid effort, but sometimes the cycle is not completed because it lacks enforcement mechanisms.'

In Mexico, evaluation and audit processes operate in a separate way. This diminishes the impact of these activities on the public sector. While coordination agencies promote evaluation within the Executive sphere, the SAO, in a separate way, provides 'congressional budget oversight [and] performance audits to assess program progress in meeting objectives and goals and generate recommendations to improve government results.'[1025]

This overview shows that the evaluation policy is still at an immature stage, regardless of the institutional efforts made in this area.[1026] The main aim of the evaluation policy has been the development of technical capacities, but its effect on policymaking is still unclear. In addition, the lack of power of the SAO has limited its role in the accountability system and it is not perceived as an effective and complementary counterpart to how the Executive evaluates itself. This discussion will become clearer through the analysis of the case studies in the next two sections.

[1023] Merino, M. 2009. "Informe sobre la evolución y el desempeño de la Auditoría Superior de la Federación." Mexico City.

[1024] Chávez, V. 2014. "Pleitos ponen en jaque al Ifai; ASF sin dientes (Lawsuits threatens IFAI; SAO toothless)." in *El Financiero* Mexico City.

[1025] Lopez-Acevedo, G., P. Krause, and K. Mackay (Eds.). 2012. *Building Better Policies. The Nuts and Bolts of Monitoring and Evaluation Systems.* Washington, D.C., US: The World Bank.. P. 176.

[1026] Cardozo, M. 2009. "La institucionalización de una cultura de la evaluación en la administración pública mexicana: Avances y desafíos pendientes." *Convergencia. Revista de Ciencias Sociales* 49:175-98.

6.2.2 The Hidden Politics of Evaluation in the Health Quality System Programme

The case of SICALIDAD delivers important evidence about the hidden politics of evaluation. Considering the context in which evaluation and audit activities take place in Mexico, looking at the case of a specific programme sheds new light to the discussion about the political rationale underpinning these instruments. This section discusses the findings of this case through the thematic framework proposed in this research (Figure 6.6).

Figure 6.6 Overview of the Hidden Politics of Evaluation in the Health Quality System Programme

Health Quality System Programme	**Purpose** *Focus on efficiency and effectiveness* *Predetermined by central government and audit body*
	Resources *Constraints to perform more sophisticated studies* *Small market of evaluators* *Data constraints*
	Process *Complex commissioning process* *Influence of commissioners*
	Outputs *Minor changes produced* *Ad hoc mechanism for the follow-up of recommendations*
	Outcomes *Diffuse effect of evaluation* *Constrained to programmatic level*

As discussed by the literature, the *purpose* of evaluation represents an important political decision because it requires conciliating the multiple interests that exist around evaluation.[1027] In the Mexican case, there is an important bias towards efficiency and efficacy which makes explicit that the 'value' of a programme is reduced to the relationship between funds spent and people benefited (inputs and outputs). SICALIDAD illustrates this; its most recent SPE provides information about its progress through indicators, i.e. coverage and budgetary performance.[1028] What arises from here is that the value given to the programme is associated mostly with managerial variables.

[1027] Rutkowski, D., and J. Sparks. 2014. "The new scalar politics of evaluation: An emerging governance role for evaluation." *Evaluation* 20(4):492-508.

[1028] National Public Health Institute. 2013a. "Specific Performance Evaluation 2012-2013 SICALIDAD." Mexico City.

This evinces a disconnection with the general objective of the programme, which is:

> To strengthen the quality of the attention in public institutions of the National Health System through the impulse of the execution and recognition of projects that help the effective access to health services.[1029]

These variables do not necessarily provide convincing evidence about the effect of a programme, but they give stakeholders a more accessible way to evaluate based on quantifiable measures, i.e. the number of beneficiaries. This is consistent with what Hedrick called 'the appropriateness of measures' in relation to the intended or unintended bias that is generated in the definition of the scope.[1030] The interesting point here is that in a context with limited space for debating evaluation, there will not be reactions to this situation even though it is a deliberate political decision of the Federal Government to prioritise the assessment of cost-effectiveness, instead of other elements like equity or transparency.

In the case of SICALIDAD, a more pertinent focus of evaluation could be the ultimate benefits (beyond perception) received by patients in terms of quality. This has been acknowledged by external evaluators who consider it necessary 'to promote research to evaluate uniformly the impact of the actions of the programme.'[1031] Nonetheless, this has not permeated the evaluation agenda. The purpose of evaluation can predetermine the results, as it implies the selection of certain kind of instrument that assesses only those variables selected by the predominant actor (coordination agencies) and that excludes the evaluated body from the process. In the opinion of a Ministry of Health official (MX-3):

> Evaluation should be based on joint decisions […] there are social programmes that should not be kept aside from evaluation, there is so much money there, so much effort, that you really need evidence of what is or not working.

[1029] Ministry of Health. 2013a. "Agreement that issues the operational rules of the Integral Health Quality System Programme for the fiscal year 2014." Mexico City.

[1030] Hedrick, T. E. 1988. "The Interaction of Politics and Evaluation." *American Journal of Evaluation* 9(3):5-14.

[1031] Universidad Autónoma Metropolitana. 2012. "External Evaluation SICALIDAD." Mexico City: Ministry of Health. P. 98.

The marginalisation of key stakeholders reduces the level of confrontation, but it also impedes evaluation from becoming a more useful tool for improvement and accountability. *Whose interests should predominate in the definition of what to evaluate and why?* The political relevance of the purpose is also associated with the generation of evidence to justify the existence or the continuity of a programme, especially when termination is perceived as a threat.[1032] Much earlier, Weiss observed that operators and managers 'are concerned not just with today's progress in achieving program goals, but with building long-term support for the programme.'[1033] A Ministry of Public Function official (MX- 2) exemplifies this:

> When there was the transition in the Federal Government from the Institutional Revolutionary Party to the National Action Party, 12 years ago, *Oportunidades*[1034] performed an evaluation with the World Bank [...] as a way to say: 'Look, our programme works' and this is said by an international organisation, not just by us.

This shows that stakeholders struggle for the predomination of their interests. Several people rely on the continuity of a programme, e.g. job security and electoral gaining. Consequently, they might have a political agenda to pursue, not always consistent among them. Evaluators might be interested in the continuity of a programme seen as an opportunity for business,[1035] while funders, i.e. international organisations, might want evidence about the effectiveness of their investment.[1036] Regarding this political agenda, an

---

[1032] DeLeon, P. 1987. "Policy Termination as a Political Phenomenon." in *The Politics of Program Evaluation*, edited by Dennis J. Palumbo. California, US: SAGE, Geva-May, I. 2004. "Riding the Wave of Opportunity: Termination in Public Policy." *Journal of Public Administration Research and Theory: J-PART* 14(3):309-33.

[1033] Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106. P. 95.

[1034] The Human Development Programme *Oportunidades* is an obligated reference in the evaluation policy in Mexico. Under the previous name of *Progresa*, this programme has been systematically and profusely evaluated (using both quantitative and qualitative approaches) since 1999 and it has set the basis for the development of evaluation capabilities, particularly in the social development sector.

[1035] For example in recommending the termination of a programme, Eddy and Berry recognise "the potential conflict that exists between those seeking to serve their clients' interests first and those who seek to serve the 'public good'." Eddy, R. M., and T. Berry. 2009. "The Evaluator's Role in Recommending Program Closure." *American Journal of Evaluation* 30(3):363-76. P. 366.

[1036] The Mexican government has signed loan contracts with the IADB for the implementation of the Results-based budget model that include the development of monitoring and evaluation systems. Source: Inter-American Development Bank. 2008. "Results-based Budget Model Support Programme-Phase I." Washington, D.C., US.

evaluator (MX-22) recalled the case of the *Piso Firme Programme*:[1037]

> When the evaluation was performed they [the evaluators] found out that providing concrete floor with antibacterial properties reduced the presence of mites and this had an impact on the incidence of gastrointestinal diseases in children, this was done by Enrique Martínez to show that the programme worked, because he aimed to be President and he wanted to disseminate the results.[1038]

The idea behind the political control of the purpose of evaluation is that it can shape the process according to the interests of stakeholders, for example, to expose the accomplishments or failures of government.[1039] This relates back to the ideas about the role given to evaluation. The establishment of a purpose related to the identification of the effectiveness of SICALIDAD generates support for the programme as it makes evident its capacity to deliver.

Another variable of the hidden politics of evaluation is the *resources*. For the performance of evaluations, SICALIDAD depends on the resources allocated by the Ministry of Health. This influences important decisions, such as the evaluators that can be hired or the instruments to be applied. As a Ministry of Health official (MX-3) recognises:

> What coordination agencies have done are evaluations based on documentary issues, there is no intentional search for evidence that can be obtained from studies of the beneficiary populations [for example] so there are a lot of obstacles to perform these exercises that you would like to do in evaluation not all programmes can be evaluated in the way you would like to, I mean, is an enormous amount of programmes and there is neither budget nor evaluators to perform those studies.

SICALIDAD clearly exemplifies this situation. In 5 years (2007–12), eight different evaluations were performed; half of those were *ad hoc* studies mostly

---

[1037] *Piso Firme* was a programme implemented by the state's government of Coahuila (a state located in the northern part of Mexico). It basically consisted in providing a concrete's floor with antibacterial properties to reduce the negative health effects on the population.

[1038] The interviewee refers to the impact evaluation leaded by Paul Gertler performed to the *Piso Firme* programme in Coahuila during 2004. Some of the most relevant findings of this study is that children under 5 years who live in households beneficiated by the programme present 18-20% less incidence of diarrhoea episodes and 20% less prevalence of anaemia than children living in non-beneficiary households. Enrique Martínez was the governor of the State of Coahuila from 1999 to 2005, this politician tried to secure PRI's nomination for the 2006 presidential election, but he was unsuccessful.

[1039] For example, during the past government period, evaluations performed to programmes such as *Oportunidades* or *Seguro Popular* were highly promoted in the media.

done by the same evaluator.[1040] This raises questions about how resources constrain the possibilities to perform more complex studies and to access a more competitive market of evaluators. With a very small pool of evaluators available, there are incentives to perpetuate this situation, particularly by those public institutions which are recurrent evaluators of federal programmes. The National Public Health Institute (INSP) has performed 50% of the SPEs to the programme and the design evaluation in 2007.[1041] Regarding this issue, an evaluator (MX-15) recognises that 'There are oligopolies, institutions where you know people are well-prepared, and it's those people who receive the evaluations, almost *de facto.*' As a result, a group of evaluators are constantly obtaining contracts from different ministries, establishing relationships that can become noxious for the independence of the process.

The availability of data required for its evaluation is also an important resource. According to a recent report about SICALIDAD, some of the data required 'was inconsistent' regarding some indicators of the programme.[1042] The literature recognises the control of data as an area of political influence.[1043] The collection of data from different sources was challenging for the evaluation of SICALIDAD, because it implied the participation of multiple stakeholders, i.e. state governments, and sometimes operators were reluctant to disclose information:

> At an operative level that happened [...] we asked for some particular data and they gave us some other thing [...] we had a case in which we asked for information in different ways and even outside the planned timescale, and they refused giving it to us [...] they were reluctant, they gave us a plain no (MX-15).

This was also observed by SAO's auditors, who rely on the information provided by the programme. According to the audit performed in 2011, the data of the programme 'does not allow evaluating the accomplishment of the general

---

[1040] For instance, the *Metropolitan Autonomous University* has evaluated SICALIDAD in three occasions since 2009.

[1041] Coordinators of the evaluations have changed but they all belong to this organisation.

[1042] National Public Health Institute. 2013f. "Specific Performance Evaluation. SICALIDAD Programme.". Mexico City. P. 21.

[1043] Bamberger, M., J. Rugh, and L. Mabry. 2012. *Real-World Evaluation: Working Under Budget, Time, Data, and Political Constraints*. California, US: SAGE, McLemore, J. R., and J. E. Neumann. 1987. "The inherently political nature of program evaluators and evaluation research." *Evaluation and Program Planning* 10(1):83-93.

objective of SICALIDAD.'[1044] This control over data can be translated into the selective provision of information (intended or not) to protect the interests of stakeholders. Omissions in the production of data systems are also an important issue, as SICALIDAD requires data from state-level health institutions that implement quality projects funded centrally. However, as evaluators identified:

> There is no integrated information system in terms of quality that allows the opportune, exact and sufficient identification of the conditions about the quality problem in health institutions in the country.[1045]

Funding for the development of evaluation studies is also a resource with political implications. As Bamberger et al. recognise, constraints in terms of data, funding, and time can affect the characteristics of an evaluation. This might favour the performance of simpler and cheaper studies.[1046] SICALIDAD has been mostly evaluated through the application of synthetic instruments based on operational information provided by the programme. This kind of evaluation is less expensive than those studies that require fieldwork.[1047] In this case, the funding restriction has even affected those mandatory evaluations. For example, the *Consistency and Results Evaluation* was cancelled in 2012 because of budgetary restrictions (MX-4).

From the perspective of the institutional framework it is interesting to observe that the obligation imposed to programmes to perform an annual evaluation might generate incentives to carry out simpler studies to comply regulation, regardless of its real contribution to the programme.[1048] The non-compliance of this rule can be seen as a resistance to be evaluated and has negative political consequences for the programme.[1049]

[1044] Supreme Audit Office. 2011. "Performance Audit. SICALIDAD.". Mexico, City.

[1045] Universidad Autónoma Metropolitana. 2012. "External Evaluation SICALIDAD." Mexico City: Ministry of Health. P. 16.

[1046] Bamberger, M., J. Rugh, and L. Mabry. 2012. *Real-World Evaluation: Working Under Budget, Time, Data, and Political Constraints*. California, US: SAGE.

[1047] The average cost of the specific performance evaluations to SICALIDAD is circa £3,500 compared to the cost of the latest *ad hoc* evaluation to the programme which cost was around £66,000. Source: Estimation based on data of the Ministry of Health.

[1048] "Social Development General Law." Mexico. Art. 78.

[1049] This has been observed recently in the resistance to evaluation found in the education sector. See: Castillo, M. A. 2015. "Gestión y evaluación: De la resistencia al desarrollo próximo." in *Segundo Congreso Latinoamericano de Medición y Evaluación Educacional*. Mexico City.

Methodological issues are also susceptible to political influence.[1050] As Chelimsky says, the political context underpinning evaluation demands giving more attention to the selection of methods, as this can be disputed by political adversaries,[1051] 'questioning the methodology is a major obstacle for evaluation to be beneficial' (MX-2). The control over the methodology can represent a mechanism for influencing the process. As described, in *ad hoc* studies evaluation units define the TORs, although evaluators usually present a technical proposal. As an evaluator of SICALIDAD (MX-15) describes: 'There is a dialogue with the commissioner to clarify some aspects mostly technical and in some cases methodological.' TORs established by the commissioner constrain evaluators from exploring any other areas. Despite this, the general opinion is that evaluators are independent enough:

> I think that they work independently and according to their own criteria, I don't believe that they are influenced [by someone], what affects the assessment is the idea that they have of the programme (MX-11).

The *process* is not exempted from political influence either. Independence is a constant element of discussion in the Mexican case. Important efforts have been made to promote the performance of external evaluations to provide an objective judgement about a programme. The key element is to what extent this independence can be preserved in a context in which the institutional framework allows excessive influence of the Federal Government and leaves a very narrow space for evaluators to make their work. This creates a scenario in which independence seems to be a core element of the discourse, but is limited in practice due to the rigidness of the evaluation policy. An evaluator of SICALIDAD (MX-15) acknowledges that he perceived 'only isolated cases of interference due to political disputes in state-level organisations [...] that altered the dynamic of the evaluation.' There do not seem to be explicit threats

---

[1050] *Oportunidades* illustrates this situation. According to an evaluator (MX-22), in the process of defining an evaluation 'there are a lot of political decisions, there is the need to evidence if a programme is doing right [...] it's not only the fact of performing evaluations, another issue is to evaluate the aspect that is under the sun, rather than the one than is in the shadow, [for example] there are very few evaluations of *Oportunidades* to assess the impact of the programme in urban localities.'

[1051] Chelimsky, E. 1995. "The Political Environment of Evaluation and What it Means for the Development of the Field: Evaluation for a New Century: A Global Perspective." *American Journal of Evaluation* 16(3):215-25.

to independence, but subtle interferences occurred as a result of the centralisation of the process and the use of predetermined frameworks. As Karlsson suggests, 'Evaluation has become a tool for control [...] and to legitimise the political decisions and priorities.'[1052]

The process concerns also commissioning evaluations. Regulation determines that ministries can contract external parties through three mechanisms: public bidding, restricted invitation, and direct adjudication.[1053] There is an exemption to public bidding when a public university will provide the service. Deadlines and procedures for this purpose are, as a generalised opinion, highly complex, and this forces the commissioner to make decisions under suboptimal conditions.[1054] An official explains: 'It is complex for us [public bidding] because of time restriction; we are always pressured by deadlines' (MX- 4). This might explain why circa 80% of the evaluations of SICALIDAD have been performed by public institutions. In addition and according to a former CONEVAL official (MX-22), some of these institutions have monopolised evaluation, producing a rise in the costs. The theory recognises that contracts in evaluation need to transcend the technical aspect and incorporate an ethical component.[1055] However, what seems more relevant to discuss are the informal rules that shape the relationship between commissioners and evaluators. If only the commissioning part is regulated there is a discretionary area in the interaction between parties. The process of evaluation in the case of SICALIDAD shows that despite the rigidity of the institutional framework, stakeholders constantly find spaces of discretion to pursue their interests.

The use of *outputs* is a matter of political concern as well. In terms of utilisation, the mechanism for the follow-up of recommendations (see Chapter 5, Section 5.1.2) is considered 'a very light instrument' (MX-8) as there is no enforcement

---

[1052] Karlsson, O. 2003. "Evaluation Politics in Europe: Trends and Tendencies." *Studies in Educational Policy and Educational Philosophy*. P. 1.

[1053] "Act on Acquistion, Leasing and Public Sector Services."Art. 26.

[1054] One of the major concerns of public agencies is that if they do not exercise public funds within the established term, the Ministry of Finance might withdraw these funds and they would no longer be able to perform the evaluation.

[1055] Bechar, S., and I. Mero-Jaffe. 2014. "Who Is Afraid of Evaluation? Ethics in Evaluation Research as a Way to Cope With Excessive Evaluation Anxiety: Insights From a Case Study." *American Journal of Evaluation* 35(3):364-76.

in the use of evaluations and it has evolved into another administrative process that ministries must follow. This produces two major concerns. First, the programme's operation and evaluation units are in charge of analysing recommendations and establishing a course of action for their implementation. This generates incentives for the selection of those considered more feasible and easy to accomplish.[1056] As a Ministry of Health official (MX-8) states:

> The follow-up system is a way of accountability, but I think that it remains too short, it makes it too easy for them [operators] [...] as far as I am concerned SICALIDAD accomplishes what it is established in the system, but still, they are not going to select the most difficult issues.

If programme operators have the freedom to select recommendations there is no real incentive to address those that can be more challenging. On the contrary, it seems to promote the fulfilment of administrative requirements with poor effects on improvement, because the focus is on the process, not on the results. To illustrate this, in its most recent work programme for the follow-up of recommendations, SICALIDAD only planned the performance of a single action: 'To develop an indicator for the matrix 2015 that provides more specific measurements about quality in services.'[1057] This reveals that the programme meets the requirements of coordination agencies, but this does not seem to have an important effect on the improvement of the programme or a substantial contribution to decision-making. The institutional framework gives coordination agencies the power to question or enforce programmes to heed recommendations, but with no real consequences in practice if they fail to comply. As one of CONEVAL's counsellors (MX-14) recognises:

> The main challenge of this mechanism is transcending from generating quality and independent evaluations to achieving impact, I think we have figured out the first part, but the second one, that we haven't accomplished it yet, [...] the follow-up process is very positive but that is something that programmes do and you can see their response to very specific recommendations, is much less the utility that evaluations have, for example, for budgetary.

A similar situation occurs in the case of the SAO performance audits. The report

---

[1056] Also, it is a recurrent argument that they cannot follow a particular recommendation because is not of their competence; they need the intervention of other actors or simply require additional funding.

[1057] Ministry of Health. 2014b. "Sistema Integral de Calidad en Salud SICALIDAD. Work Programme." Mexico City.

issued in 2011 contained recommendations mostly related to the lack of indicators and mechanisms to assess its accomplishment. The audited entity is expected to deliver promptly the information requested, to cooperate with the auditing process, and to respond to the observations made by the SAO; yet they are only expected to establish actions of improvement or to provide arguments for not heeding recommendations.[1058] Although utilisation is contemplated in the institutional framework, in practice it is a weak instrument that does not clearly promote the use of outputs, mostly because as an auditor (MX-12) states, the SAO "lacks 'teeth' to enforce public agencies to act in a certain way."

Utilisation as a core element of the theory developed around evaluation suggests that if institutions are not clearly being benefited from this activity, 'there is little justification for investing the resources necessary to conduct an evaluation.'[1059] It is difficult to determine an accurate rate of return from evaluation and auditing by only considering instrumental benefits when political returns from these activities cannot be quantified but are also quite important.

Regarding dissemination, regulation enforces the publication of all compulsory evaluations in the Official Federation Journal[1060] and in the ministries' websites.[1061] This reduces the space of discretion for stakeholders to selectively disseminate outputs. However, this does not mean that their interests cannot be affected. A SICALIDAD operator explains that an evaluator disseminated the results of a study without the programme's authorisation. The findings were not positive because, according to this interviewee, he 'evaluated according to his models and not considering the context of the programme, so this harmed the programme considerably' (MX-17).

This reveals two important issues. On the one hand, getting out of the predetermined framework for evaluating the programme generates conflicts

---

[1058] "Federal Accountability Law." Art. 32.
[1059] Neuman, A., N. Shahor, I. Shina, A. Sarid, and Z. Saar. 2013. "Evaluation utilization research—Developing a theory and putting it to use." *Evaluation and Program Planning* 36(1):64-70. P. 64.
[1060] "Social Development General Law." Mexico. Art. 79.
[1061] "General guidelines for the evaluation of the federal programmes of the Federal Public Administration." Mexico. Title III, chapter VII.

between stakeholders, particularly if applying alternative models deriving in negative findings. On the other hand, the dissemination of negative aspects of the programme is considered a threat because it exposes operational deficiencies that can affect them individually, e.g. losing their job. Even though regulation enforces dissemination, this does not completely limit the potential influence of stakeholders to align outputs to their interests and 'to present themselves in a positive light.'[1062] For example, the award promoted by the Ministry of Public Function 'to identify the programme's commitment to heed evaluation recommendations'[1063] was an interesting opportunity for SICALIDAD to promote its positive aspects. This prize reflected its 'remarkable appropriation of the logic of evaluation results to generate internal improvement processes'[1064] and it allowed the programme to disseminate those evaluation outputs that clearly justify its existence, particularly in a political moment of transition in which many programmes were at risk of termination.[1065] This was a mechanism of dissemination that legitimated the programme through the intervention of the external reviewers that awarded the prize.

Finally, in terms of *outcomes*, SICALIDAD provides some interesting lessons about the effect of evaluation in the improvement of the programme. Several interviewees were asked if they could recall a substantial change derived from the findings of an evaluation. Most of them agreed on saying that only minor adjustments have been made. Long-term effects of evaluation are not easy to define because of the complexity of isolating them from other variables that might have intervened. Evaluation can find resistance because, as Picciotto highlights:

[1062] O'Brien, T., S. Payne, M. Nolan, and C. Ingleton. 2010. "Unpacking the Politics of Evaluation: A Dramaturgical Analysis." *Evaluation* 16(4):431-44. P. 432.

[1063] Cejudo, G. 2011. "De las recomendaciones a las acciones: El uso de la evaluación en la mejora del desempeño del gobierno." in *De las recomendaciones a las acciones: La experiencia del Premio 2011 Programas Federales Comprometidos con el Proceso de Evaluación*, edited by G. Cejudo and C. Maldonado. Mexico City: CIDE-SFP. P. 8.

[1064] Maldonado, C. 2011. "El programa SICALIDAD: Ejemplo de gestión pública anclada en la evaluación para la mejora continua." in *De las recomendaciones a las acciones: La experiencia del Premio 2011. Programas Federales Comprometidos con el Proceso de Evaluación*, edited by G. Cejudo and C. Maldonado. Mexico City: CIDE-SFP. P. 67.

[1065] Presidential elections were held in July 2012 and the elected government took office in December of that year.

> The impetus [...] has often come from parliaments, supreme audit organizations, budget authorities or funding agencies. Internally, there are fears that budgetary allocations may be reduced as a result of evaluations; reluctance to share control over information flows; distrust about performance-based evaluations.[1066]

The fact that evaluation is already an institutionalised practice that has become a part of the programme's routine is an important progress, but other elements still need to be addressed. The first one is the quality of the studies, as currently its performance does not guarantee that studies are being properly conducted.[1067] The other is the promotion of more effective mechanisms for the utilisation of findings. In the case of SICALIDAD, evaluation and auditing can represent core elements of its operation as the problem that is tackling the programme is complex and multifactorial. In Mexico, quality of health services 'is heterogeneous and its efficiency is reduced, causing the prevalence of dissatisfaction in users and providers.'[1068] Consequently, the production and utilisation of evidence is crucial to identify the most critical areas to address.

Despite the efforts made by the Federal Government, it appears that evaluation is not completely embedded into policymaking, even less in the political sphere. After the 2012 election, SICALIDAD suffered minor adjustments to its operation (transiting from 'programme' to 'strategy'), but it has not reached an important level of notoriety in the political sphere. For example, President Peña Nieto recently declared: 'Let's have an area within the health sector specifically in charge of monitoring the quality of services provided by the institutions of the National Health System'.[1069] This suggests that politicians are not entirely aware of the work carried out by ministries or of the existence of findings that show the progress made in this area. So political changes that might interfere with policy continuity of these changes are seen as an opportunity to start from

---

[1066] Picciotto, R. 1999. "Towards an Economics of Evaluation." *Evaluation* 5(1):7-22. P. 14.

[1067] The quality issue in evaluation has fostered, for example, the discussion about the pertinence of the certification of evaluators. See: Davies, R., D. Randall, and R. E. West. 2015. "Using Open Badges to Certify Practicing Evaluators." *American Journal of Evaluation* 36(2):151-63.

[1068] Alcántara, M. A. (Ed.). 2012. *La calidad de la atención a la salud en México a través de sus instituciones: 12 años de experiencia*. Mexico City: Ministry of Health. P. 12.

[1069] Rodríguez García, A. 2014. "Peña commands the creation of an area to overview quality of health services." in *Revista Proceso*. Mexico City.

scratch.[1070]

SICALIDAD allows looking at the consequences derived from a very rigid and centralist evaluation policy, a weak external audit body and an embryonic market of evaluators. These variables have generated spaces of discretion that stakeholders use in the pursuit of their interests. Moreover, the lack of participation from external stakeholders (NGOs, charities, media) in questioning the results of audits and evaluations and the process itself leaves these activities in a position of having very little influence. The next section focuses on the case of the Health Caravans Programme, which also provides consistent findings about the hidden politics of evaluation in Mexico.

[1070] See: Wilson, C. A. 2013. *Public Policy: Continuity and Change*. USA: Waveland Press.

6.2.3 The Hidden Politics of Evaluation in the Health Caravans Programme

The case of the Health Caravans Programme is also analysed through the lenses of the thematic framework proposed in this research. The analysis aims to present additional evidence about how politics affects evaluation, considering the Mexican institutional framework. Due to its nature, Health Caravans has had more notoriety in the public sphere. Local media, for example, often spread the word about the arrival of a caravan to a particular location.[1071] It is also a more salient programme in terms of budget (circa £30 million in 2012).[1072] These particularities give this case a different perspective to the hidden politics of evaluation (Figure 6.7).

Figure 6.7 Overview of the Hidden Politics of Evaluation in the Health Caravans Programme

Health Caravans Programme	**Purpose** *Predetermined by central government* *Limited scope to a range of variables*
	Resources *Small market of evaluators* *Asymmetry of information for commissioners*
	Process *Discrepancies between stakeholders about the evaluation process* *Closeness between commissioners and evaluators*
	Outputs *Use constrained to improvements in operation* *Minor political use of evaluation outputs in the media*
	Outcomes *Diffuse effect of evaluation* *Political sustainability*

In terms of the *purpose*, coordination agencies control the periodicity, methodology, and scope of evaluation. The literature recognises these as political decisions because this stage imposes 'the challenge of reconciling a wide range of understandings of, and purposes for, program evaluation among these stakeholders.'[1073] A power struggle arises between coordination agencies in relation to the definition of the programmes to evaluate and the kind of evaluation to apply. According to members of the Ministry of Finance,

---

[1071] See: Estrada García, K. 2015. "Health Caravans come to Ascensión." in *Diario.mx*. Chihuahua, Mexico, Trejo, Á. 2015. "Health Caravans arrive to Epazoyucan." in *Criterio*. Hidalgo, Mexico.

[1072] Source: National Public Health Institute. 2013b. "Specific Performance Evaluation. Health Caravans Programme 2012-2013.". Mexico City.

[1073] Rutkowski, D., and J. Sparks. 2014. "The new scalar politics of evaluation: An emerging governance role for evaluation." *Evaluation* 20(4):492-508. P. 493.

coordination agencies receive a list of evaluation needs from programmes that are negotiated for the final definition of the AEP (MX-18-22). This exposes an area of discretion produced by the lack of explicit criteria for the selection of programmes and evaluation instruments. While programmes like Health Caravans have been systematically evaluated, others have just recently been incorporated into this process, e.g. programmes from the Ministry of Tourism.

The defined purpose of evaluation is limited mostly to effectiveness and operational issues, leaving aside other features that might generate evidence of a different nature. For example, the purpose of Health Caravans' most recent evaluation was:

> To describe and analyse the operation of Health Caravans in order to detect opportunity areas and strengths that allow the development of a continuous improvement model of its macro processes to achieve the goals and targets of the programme in a more efficient, effective and opportune way.[1074]

In the opinion of an academic (UK-1), the narrowness of the purpose of evaluation to determine effectiveness shows:

> A bias to grant *a priori* the public value generated by the programme [...] it looks only at the accomplishments of the programme but there is no previous questioning of the relevance, the public value and the mere existence of the programme.

This gives space for discussing an important issue of the Mexican case. Here evaluation relies on the assumption that the policy is appropriate to address the problem. Constraining this activity to the discussion of the effectiveness of programmes can help explaining why evaluation has not penetrated the political discussion. For instance, the provision of health services in remote locations by Health Caravans seems a palliative to a much more complex problem: access. Therefore, the study of variables beyond the effectiveness of the programme and its operational issues is required for obtaining evidence about better ways for tackling the problem. According to Juárez-Ramírez et al., Mexico has a serious problem of inequality in the access to health services, especially among vulnerable groups of the population such as the elderly and

---

[1074] Universidad Intercultural del Estado de México. 2013. "Health Caravans Evaluation Programme in High and Very Highly Marginalised Locations ". Mexico City. P. 1.

migrants.[1075] Even though the focus of the evaluations to Health Caravans has been its effectiveness as an alternative for the provision of health services in remote locations, its lack of impact due to the magnitude of the problem has been acknowledged, as 'it is insufficient considering the real health needs of the population that it serves.'[1076]

This resonates with Clarke's ideas about success. He states that evaluation contributes to the identification of 'winners and losers.'[1077] Consequently, the variables that transmit the accomplishment of targets are a good mechanism to demonstrate the 'value' of a programme. In the case of Health Caravans, evaluations have highlighted its accomplishments in terms of coverage and trained staff.[1078] However, these indicators provide little information about its capacity to reduce inequalities.

Resources represent another variable susceptible of political influence. The centralisation of financial resources for evaluation leaves ministries in a position of mere executors of decisions taken by others. Financial resources predetermine not only the number of evaluations to perform during a fiscal year, but also its nature, as some studies are more expensive than others.[1079] For example, the cost of the most recent *ad hoc* evaluation to Health Caravans was approximately £431,000 (nearly 1.2% of its 2014 budget[1080]). Considering operational costs it is not possible to allocate many more resources to evaluation; *ergo,* there are constraints beyond the control of the programme, as political will and commitment to evaluation are not the only elements required

---

[1075] Juárez-Ramírez, C., M. Márquez-Serrano, N. Salgado de Snyder, B. E. Pelcastre-Villafuerte, M. G. Ruelas-González, and H. Reyes-Morales. 2014. "La desigualdad en salud de grupos vulnerables de México: Adultos mayores, indígenas y migrantes." *Revista Panamericana de Salud Pública* 35:284-90.

[1076] National Public Health Institute. 2013b. "Specific Performance Evaluation. Health Caravans Programme 2012-2013.". Mexico City. P. 2.

[1077] Clarke, J. 2008. "Performance Paradoxes: The Politics of Evaluation in Public Services." Pp. 120-34 in *Public Services Inspection in the UK. Research Highlight in Social Work*, edited by H. Davis and S. Martin. London, UK: Jessica Kingsley Publishers.

[1078] National Public Health Institute. 2013b. "Specific Performance Evaluation. Health Caravans Programme 2012-2013.". Mexico City.

[1079] As an example, the Health Caravans Programme carried out in 2011 a Specific Performance Evaluation which cost was 65, 384.23 Mexican pesos (approximately 3,500 GBP) and a complementary evaluation that cost 5, 490,000 Mexican pesos (circa 293,000 GBP), with these examples we can observe the range that exists in the cost of evaluations to public programmes.

[1080] Estimate based on information from the Ministries of Health and Finance.

for this activity to succeed.

This also shows the unfeasibility of evaluation of all existent programmes, despite the proportion of their cost in the overall budget of the Federal Government[1081] and the selection of instruments suitable to the needs of the programme. It is difficult to carry out more sophisticated and possibly more expensive studies. This might also imply sacrificing quality and prioritising the cost variable. As a Ministry of Health official (MX-5) recognises, 'Often quality is sacrificed for cost […] when criteria [for evaluation] was not properly established.' This issue has been recently discussed in the academic sector:

> The evaluation market is still incipient: lots of evaluators are relatively new in the evaluation field, there are not enough signals about the quality of evaluations, nor an instance that qualifies the solidness of an evaluator […] In addition the acquisition processes in the Federal Government are good for identifying good prices, but not necessarily good quality.[1082]

In the case of Health Caravans, operators were very enthusiastic about performing more complex studies, but they are limited by the resources available. Most evaluations have been desk-based studies that have provided information for internal consumption (MX-11). The evaluations commissioned and funded by CONEVAL are standardised studies that do not entirely meet the programme's needs. This generates a political dispute because 'it involves power relations and decisions about rules and resources, particularly the allocation of funding.'[1083]

Like SICALIDAD, the limited pool of evaluators has also been a weak link in the evaluation system. It is important to recall that evaluation bloomed in Mexico after the results-based Management model was institutionalised in 2007, generating an important increase in the demand for evaluation studies. As an academic explains (MX-1): 'There wasn't an installed capacity [of evaluators],

---

[1081] According to information of the Synthetic Model of Performance developed by the Ministry of Finance, in 2012 were identified 1022 budgetary programmes operating in the Federal Government.

[1082] Cejudo, G. 2011. "De las recomendaciones a las acciones: El uso de la evaluación en la mejora del desempeño del gobierno." in *De las recomendaciones a las acciones: La experiencia del Premio 2011 Programas Federales Comprometidos con el Proceso de Evaluación*, edited by G. Cejudo and C. Maldonado. Mexico City: CIDE-SFP..

[1083] O'Brien, T., S. Payne, M. Nolan, and C. Ingleton. 2010. "Unpacking the Politics of Evaluation: A Dramaturgical Analysis." *Evaluation* 16(4):431-44. P. 433.

instead it has been created since; universities have started to direct researchers toward evaluation; consultancies have been created.'

The offer of evaluators is insufficient for the existent demand.[1084] Consequently, a reduced group of evaluators is constantly hired due to the lack of options.[1085] Some of the evaluators of SICALIDAD have also evaluated Health Caravans, i.e. National Public Health Institute, and other programmes operated by the Ministry of Health.[1086] This also occurs in the case of *ad hoc* evaluations, when the National Autonomous University of Mexico has performed 42% of the studies.[1087] This reveals another problem in terms of the capacity to perform evaluation studies. A former CONEVAL official (MX-24) explains:

> At first we didn't have enough knowledge about evaluators […], so we turned to universities and this process was not easy, they were used to write papers but not to write evaluation reports.

The development of evaluation capabilities has been an incremental process but it has not favoured the expansion of the market. The limited group of evaluators that exists have gained experience and knowledge about the operation of the evaluation policy, which gives them a competitive advantage regarding those academicians that have the technical expertise to evaluate but who have not entered into this privileged group. According to Maldonado, both the legal complexity for commissioning evaluations and the prevalence of public institutions in the market of evaluation 'induces a remarkable barrier of entrance to new, private and individual consultant services providers and puts a bunch of academic institutions in an oligopoly of the market of evaluation.'[1088]

[1084] To illustrate the dimension of this situation, the AEP 2013 established the evaluation of 210 federal programmes for this year. These programmes are operated by different agencies of the Federal Government and require the participation of multiple evaluators.

[1085] As an example, the SPE coordinated by CONEVAL in three different periods of time (2008-2009, 2009-2010 and 2011-2012) shows that approximately 32% of the evaluators in charge of conducting the evaluation have been hired at least in two occasions out of three in which this evaluation has been performed. Due to the numerous evaluations that needed to be hired, it is not possible to identify specialists for every subject; instead, evaluators participate in the assessment of several programmes that are linked to their broadest expertise.

[1086] For instance, most evaluators of Health Caravans also have performed studies for the *Seguro Popular* programme, one of the most salient health initiatives of the Federal Government.

[1087] This also occurs with the Metropolitan Autonomous University in the case of SICALIDAD.

[1088] Maldonado, C. 2012. "Más allá de las instituciones: La evaluación y la gestión por resultados en México. Una perspectiva de mercado." in *XVII Congreso Internacional del CLAD sobre la*

A Ministry of Health official (MX-4) justifies commissioning evaluations to recurrent institutions: 'They have the infrastructure and knowledge to perform this kind of evaluation. They even take blood samples, which require a lot of resources.' This resonates with the lack of evaluation capabilities in the country. As Rist et al. state, human capital is an element of evaluation capacity. 'There needs to be good capacity in terms of the quantity and quality of M&E human resources, both within the organization (M&E staff) and outside the organization (external evaluators).'[1089]

This situation produces a lack of market signals about quality, which according to a CONEVAL official (MX-9) 'has been very heterogeneous.' There are no mechanisms for commissioners to discriminate among evaluators and 'a consultancy firm can disappear and reappear under a new name to make systematically poor evaluations and survive in the market because there are no means to eliminate them'(UK-1). In line with this, an NGO reflects about the role of coordination agencies:

> They need to go further overviewing the compliance of the regulation. They must have a proactive role in the provision of data and evidence to commissioners about the quality and costs of evaluations, reducing information asymmetries.[1090]

This is observable in *ad hoc* evaluations. SICALIDAD, Health Caravans, and *Seguro Popular* have consistently commissioned this kind of study. For the 2009–12 period, the average cost of each one of these evaluations was circa £170,000 for a 3-month contract.[1091] However, the scope, methodology, and complexity of the studies differ. For example, the 2012 evaluation to Health Caravans was an analysis of the evaluations performed during 2007–11[1092] (meta evaluation). In contrast, the 2010 evaluation to *Seguro Popular* comprised the study of two groups of women who had been treated for breast cancer at

---

Reforma del Estado y de la Administración Pública. Cartagena, Colombia. P.8.

[1089] Rist, R. C., M. H. Boily, and F. Martin. 2011. *Influencing Change. Building Evaluation Capacity to Strengthen Governance*. Washington, D.C., US: The International Bank for Reconstruction and Development/The World Bank. P. 7.

[1090] GESOC, A.C. 2009. "Monitoreo de la Confiabilidad de las Evaluaciones a Programas Federales (2007-2008)." P. 8.

[1091] Estimation based on information of the Ministry of Health.

[1092] Universidad Nacional Autónoma de México. 2012. "External Evaluation. Health Caravans Programme.". Mexico City.

selected hospitals.[1093] Consequently, there are no criteria to determine a range of costs derived from the characteristics of the study (complexity, length, etc.). The lack of options gives evaluators the prerogative to set the price or at least to be in a better position for negotiating.

The *process* of evaluation has also political considerations. Discretionary spaces can be found in the selection of methodology, during the commission of evaluations and in the interaction among stakeholders. A clear example is the complexity of the commissioning process, as an interviewee (MX-24) suggests 'there are perverse incentives for public entities [...] to avoid public bidding for simplifying the administrative work of finding people to carry out the studies.' This keeps the market reduced and promotes a closer relationship between commissioners and evaluators in universities, particularly those publicly funded:

> There is a very serious reputation game with universities [...] I think that the best evaluations are done within academic institutions, but there are also very poor evaluations that rely on the institution's reputation (UK-1).

This creates incentives for preserving the *status quo.*[1094] Evaluation represents an important area of business for researchers while at the same time it also simplifies the work of commissioners. The SPE illustrates this as there are recurrent evaluators in different years and programmes of the Ministry of Health.[1095] Some interviewees identified that evaluators who are familiar with the administrative process can take advantage of this situation through informal arrangements. As an evaluator (MX-22) points out:

> It can be a very flexible situation, if you want to perform an evaluation but you belong to a private university you go to a public institution, you pay the overhead, the contract is established under the name of a researcher from that university [...] you perform the evaluation and get paid.[1096]

[1093] National Public Health Institute. 2012. "Evaluation of the effects of the catastrophic expenses fund in breast cancer patients ". Mexico City.
[1094] These universities and centres receive public funds for their operation. In the case of public research centres, these belong to the Federal Public Parastatal Administration, so they are part of a scheme in which public programmes are evaluated by researchers that belong to public universities.
[1095] Source: Specific Performance Evaluations: Health Caravans, SICALIDAD and *Seguro Popular* 2008-2012.
[1096] This situation has also led researchers at private institutions to consider that they are in

A Ministry of Health official (MX-6) explains that the familiarity of the evaluators with the nuances of the process can also impact on the quality of the studies. 'In the commercial relationship that is established to perform an evaluation the evaluator becomes too confident knowing that once a project is authorized, the flow of resources is guaranteed.' The closeness of the market generates space for informal agreements that can compromise independence, as House observes in the 'socialization' between evaluators and commissioners.[1097] This can represent a threat to independence...

> Evaluators need freedom to plan and implement the particular study design that will bring the best answer to the question asked, and they need freedom to determine their findings and report on what they've found.[1098]

As both commissioners and evaluators have interests at stake, constant interaction can lead to the establishment of informal rules around the process of evaluation, e.g. in the provision of data. According to a Ministry of Health official (MX-8), 'Programmes are given the opportunity to update documents', which will be part of the evaluation. In the case of the SAO audits performed to Health Caravans in different states of the country, for example, it was reported that some operative documents were not provided to auditors until later in the process.[1099]

Once a contract is assigned, the process of evaluation continues to a stage of monitoring and supervision. At this point, independence might be compromised due to the interference of stakeholders. In the Mexican case some evaluators recognised that they have perceived the influence of others, mostly evaluation units and programme operators. However, they sustained that this does not have the political intention of manipulating the outputs:

> I think that it is in a good sense, they [public agencies] know well their programmes and they alert about some misinterpretations [...] giving us a

disadvantage to participate in evaluation processes for the Federal Government, as they have to deal with several administrative processes that might discourage public agencies to hire them.

[1097] House, E. R. 1997. "Evaluation in the Government Marketplace." *American Journal of Evaluation* 18(1):37-48.

[1098] Chelimsky, E. 2009. "Integrating Evaluation Units into the Political Environment of Government: The Role of Evaluation Policy." *New Directions for Evaluation* 123:51-66. P. 55.

[1099] Supreme Audit Office. 2013b. "Financial Audit. Federal resources transfer through the agreement between the Ministry of Health and the State of Zacatecas." Mexico City. P. 2.

dose of reality from their experience, but never with the aim of modifying the sense of the evaluation or the content of recommendations (MX-1).

This finds echo in the discussion about the agendas of stakeholders. In Health Caravans, conflicts have arisen between operators and evaluators. An operator (MX-11) describes that they disagreed with an evaluator due to the approach taken: 'They had a different vision, they [...] only evaluated a particular moment in different locations, we considered that the study was not valuable to us.' From this statement it is possible to observe that despite the inconformity of the programme, the findings of the evaluation were not changed. This reveals that not always in a relationship between operators and evaluators is the latter in a more disadvantaged position. The institutional framework in Mexico, for instance, withdraws the right of commissioning from operators and gives it to a third party (coordination agencies or evaluation units). Without this power, perhaps the relationship between evaluators and operators can become more equitable as evaluators do not perceive operators as a direct client. Nonetheless, it is an important issue for evaluators to be politically sensitive during the process of evaluation, mostly in the way they communicate findings, as an evaluator (MX-10) expresses:

> You want to make an excellent job, but this does not guarantee you per se that you will be hired again, this does not necessarily mean that evaluators might become co-opted or compromised to say things that are not true, but maybe to be careful in the way they say it.

There might be incentives for evaluators to please commissioners as evaluation is a business opportunity. Evaluators cannot openly accept this situation, but preserving a good relationship with commissioners for future contracts seems a logical behaviour. Bechar and Mero-Jaffe observed this in the evaluation of an educational training programme, about which they described the difficulties that emerged with the programme's head:

> He did not comment on positive evaluation findings, while he rejected negative findings and all recommendations, instead offering alternative interpretations or remarking that the issue was already taken care of.[1100]

[1100] Bechar, S., and I. Mero-Jaffe. 2014. "Who Is Afraid of Evaluation? Ethics in Evaluation Research as a Way to Cope With Excessive Evaluation Anxiety: Insights From a Case Study." *American Journal of Evaluation* 35(3):364-76. P. 368.

A recurrent topic of discussion between operators and evaluators has been the definition of the potential population that Health Caravans must serve. An operator (MX-11) stated that the quantification of people who lack access to health services is not the responsibility of the programme: 'It does not concern us to determine coverage in a particular location', but this was an element highlighted in the last *Consistency and Results Evaluation*,[1101] and also mentioned in an evaluation of 2011.[1102] Regarding this the programme stated:

> The objective population of the programme is focalized and it corresponds to the population that inhabits locations that formed part of the routes […] they do not receive attention from other health units.[1103]

This challenges the role of evaluators as advocates. Mohan clearly points out the complexity of coping with the interests vested in evaluation:

> Being responsive only to the evaluation sponsor or to a select few stakeholders could be perceived by others as a bias or taking sides. Such perceptions could adversely affect the credibility of the evaluation and could slowly over time erode the credibility of the evaluator and the evaluation organization. In the public policy environment, credibility comes, in part, from clearly demonstrating the independence of evaluators and evaluation organizations from sponsors and stakeholders.[1104]

This matters because of the challenges that evaluators face to preserve independence and credibility and it leads to reflecting about the role of the other stakeholders, e.g. sponsors, operators.[1105] The way in which the Mexican institutional framework distributes power generates a different dynamic with which coordination agencies and evaluators appear to be in a more favourable position than operators and evaluation units.

The *outputs* of evaluation in the Health Caravans case are also an interesting

---

[1101] Ministry of Health. 2011-2012. "Consistency and Results Evaluation. Health Caravans Programme.". Mexico City.
[1102] Martín del Campo Rodríguez, A. C. 2010-2011. "Specific Performance Evaluation. Health Caravans Programme.". Mexico City.
[1103] Ministry of Health. 2011b. "Specific Performance Evaluation 2010-2011. Health Caravans Programme. Institutional Statement.". Mexico City.
[1104] Mohan, R. 2014. "Evaluator Advocacy: It Is All in a Day's Work." *American Journal of Evaluation* 35(3):1-7. P. 398.
[1105] See: Datta, L. E. 1999. "The ethics of evaluation neutrality and advocacy." *New Directions for Evaluation* 1999(82):77-88, Greene, J. C. 1997. "Evaluation as Advocacy." *American Journal of Evaluation* 18(1):25-35, House, E. R., and K. R. Howe. 1998. "The Issue of Advocacy in Evaluations." *The American Journal of Evaluation* 19(2):233-36.

area of analysis. As described, use is promoted through weak instruments that establish a process for the follow-up of recommendations with little impact in the overall policymaking process. The follow-up of recommendations made by Health Caravans has mostly been constrained to operational issues. For example:

> To select indicators reflecting changes in the beneficiaries' health.
> To identify operation areas not currently overseen.[1106]

Derived from this, the programme developed some technical tools to improve operation. In 2011, CONEVAL recognised it for 'the development, systematisation, and operation of an information system for the control and identification of beneficiaries.'[1107] Another key issue is the political use of findings linked to the notoriety of Health Caravans and the attention given by the media. For instance, it is common to observe local politicians promoting the achievements of the programme:

> The programme in 2014 undertook 101,688 health prevention and promotion actions through 25,600 consultations in 243 rural communities that lack access to permanent health services.[1108] *Ministry of Health and Social Well-being, State of Colima.*

> As of today with Health Caravans we are going to offer medical services and epidemiological surveillance to those locations that lack or have limited access to traditional health services. 61 health caravans will go to the most remote sites of our state with prevention and promotion activities and ambulatory attention for circa 90, 000 people.[1109] *Manuel Velasco, Governor of the State of Chiapas.*

> Health Caravans has provided 43,652 general consultations, including prenatal control, nutritional state of children under 5, hypertension, diabetes, cervical and breast cancer detection, among other services to benefit the most needed.[1110] *Coordinator of Health Caravans in the State of Veracruz.*

These statements highlight two relevant points. On the one hand, it is

---

[1106] Ministry of Health. 2014a. "Follow-up mechanism of recommendations. Work plan. Health Caravans Programme.".
[1107] National Council for the Evaluation of the Social Development Policy. 2011a. "II Award "Good practices in the use of monitoring and evaluation outputs in policymaking" 2010-2011."
[1108] 2015. "Health Caravans provides 25,600 health consultations." El Correo de Manzanillo
[1109] Government of the State of Chiapas. 2013. "We take health where it is needed: Manuel Velasco." Chiapas, Mexico.
[1110] *2013a. "Health Caravans registers more than 40, 000 consultations in Veracruz." in La Crónica. Mexico City.*

interesting that the focus given by politicians to those indicators of the programme are quantifiable, e.g. the number of consultations. This is more politically profitable because it transmits in a succinct way the idea of success, understood as people benefited and locations were served, and also it shows where the money has been spent. On the other hand, all this positive recognition comes from politicians from the same party in office in the Federal Government.[1111]

At the national level, the political use of outputs has been observed in programmes politically relevant for the government and that have received more attention from the media.[1112] After the *Consistency and Results Evaluation* 2011–12 was published, a national newspaper published an article entitled *CONEVAL Evidences Health Caravans' Flaws,* which focused again on the deficiencies related to the definition of the target population, the lack of a registry of beneficiaries and instruments to assess its impact on the population's health. According to the programme's operators (MX-11):

> There have been some negative opinions […] that state that the programme does not assess the health impact, but it does, only that this information was not requested by the evaluator. As a consequence of this press article, the interviewee added it affected us politically and in terms of budget […] an 80 million pesos reduction was made.

This was observed in a reduction of the programme's budget in 2011.[1113] However, this was prior to the publication of the report, so it suggests that other factors might have intervened. As the research of Bamberger suggests, data 'is also used to help with future investment planning and budgetary allocations.'[1114] For example, in the context of the budgetary process, CONEVAL provides annual information to deputies about social programmes based on evaluation outputs, as there is a need 'to prioritize the access to health services

---

[1111] Governors from the states of Chiapas, Veracruz and Oaxaca all belong to the Institutional Revolutionary Party (PRI).

[1112] There can be identified two federal programmes that have received plenty of attention of the media: *Seguro Popular* and *Oportunidades*. Both were considered iconic of the 2007-2012 Federal Government Administration. *Oportunidades*, for example, in 2012 had a budget of 66,132,502,318 Mexican pesos (approximately $3,547,000,000$ GPB).

[1113] National Public Health Institute. 2013b. "Specific Performance Evaluation. Health Caravans Programme 2012-2013.". Mexico City. P. 8.

[1114] Bamberger, M. 1991. "The Politics of Policy Evaluation in Developing Countries." *Evaluation and Program Planning* 14(4):325-39. P. 334.

to vulnerable groups (people living in rural zones, indigenous, disabled people, women) and to emphasise actions that allow effective access to health services'[1115], which is clearly linked to the focus of Health Caravans. Nonetheless, as a former deputy (MX-23) recognises, 'Evaluations are rarely used [...] I would even say that the reports of the SAO are more commonly used [...] the SAO is independent and more related to deputies.' The director of an NGO (MX-16) also acknowledges that budget allocation relies more on political than on technical knowledge:

> There will always be a political component, the discussion and approval of budget takes place in the most political space existent which is Congress [...] but that should not be the only variable, it is necessary to have other variables, and in those extreme cases where there is systematic evidence that a programme has not improved, they should make a decision.

The use of outputs in this case appears to be mostly limited to the public recognition of achievements for political or electoral purposes and minor use for the improvement of the programme and budgetary decisions. Regarding the information produced by the SAO about Health Caravans, it is important to acknowledge that this organisation has audited the programme as a consequence of the financial audit of funds transferred from the Federal Government to states for the operation of the programme. In this context, the identification of misuse of funds has been an element that has reached notoriety. In the case of the State of Morelos, for instance, the SAO detected 'non-compliance of the regulation, fundamentally in payments for personal services of the Health Caravans programme.' Nonetheless, this seems to be an isolated incident.[1116]

In terms of dissemination, like SICALIDAD Health Caravans is enforced to make public the evaluation reports and information about contracts, so there is no space for manipulation of findings. Nonetheless, according to a CONEVAL's study, the Ministry of Health discloses selectively information about the evaluations of its programmes, as not in all cases is complete information

[1115] National Council for the Evaluation of the Social Development Policy. 2014b. "Considerations for the 2015 Budget Process." Mexico City.
[1116] Supreme Audit Office. 2013a. "Financial Audit. Federal resources transfer through the agreement between the Ministry of Health and the State of Morelos." Mexico City. P. 5.

published.[1117] Regarding the *ad hoc* evaluation performed in 2012 on Health Caravans, there is no information about the cost, the length of the contract, and the mechanism used for its allocation.[1118] Moreover, in the index of dissemination developed by CONEVAL, the Ministry of Health is below the average (71.43) due to the cost of the evaluation and data of the evaluator, some of the variables in which the level of compliance is lower.[1119] There does not seem to be any relevant consequence associated to the lack of compliance of the rules, except for the 'symbolic sanction' of making this information public.

The dissemination of audit findings is diluted within the wide range of information disclosed through the *Annual Accounts Report* and individual reports related to the transfer of federal funds to states.[1120] Outputs are rarely mentioned by the media or other stakeholders, except for those cases in which an important finding stands out, as in the case of the State of Coahuila, when the SAO found out that 'only 36% of the funding allocated for the operation of Health Caravans was used';[1121] *ergo,* dissemination of outputs was minor. This resonates with the work of Lawrenz et al. about the implementation of dissemination techniques to promote use. For them: 'If the evaluator intends for the report to be used, report dissemination should be tailored for each audience in terms of scope, sequence, timing, and presentation format.'[1122] In the case of the audits to Health Caravans, information is buried into much data about other programmes and strategies operated with federal resources. As a result, it is difficult to access this information.

Finally, in relation to the *outcomes* of evaluation, observing the net effects of

---

[1117] The Guidelines for Evaluation that regulate this activity within the Federal Government state the information that public organisations need to make public in relation to the evaluations performed (See Chapter 5).

[1118] This also occurs in the case of the compulsory evaluation performed in 2011, where the evaluation unit in the Ministry of Health did not make public the data about the evaluation (e.g. cost), what is more, it was only published an executive summary of the evaluation with no data from the evaluator who performed the study. Ministry of Health. 2013b. "Health Caravans Evaluations."

[1119] National Council for the Evaluation of the Social Development Policy. 2012b. "Dissemination Index of External Evaluations to Social Development Federal Programmes." Pp. 9-10.

[1120] Supreme Audit Office. 2014b. "General Public Account Report 2013." Mexico City.

[1121] Pérez, L. 2015. "Coahuila spent only 36% of the Health Caravans funding." in *El Siglo*. Coahuila, Mexico.

[1122] Lawrenz, F., A. Gullickson, and S. Toal. 2007. "Dissemination. Handmaiden to Evaluation Use." *American Journal of Evaluation* 28(3):275-89. P. 287.

evaluation on policymaking, beyond some minor operational changes, is a complex task. Evaluation has not received enough attention outside government and there is no active debate about the utility of evaluation and the quality of the reports, except for some NGOs, i.e. *GESOC, México Evalúa*.[1123] Long-term effects of evaluation and audit processes of Health Caravans cannot be clearly identified. Perhaps the most obvious effect has been the continuity of the programme after political changes took place in 2012. There were no substantial modifications to its operation and there was a 25% increase to its budget in the federal budget project of 2015 (in relation to 2014).[1124] This suggests that the programme is still part of the current government's agenda, regardless of its real contribution to the problem of inequality in the access to health services.

This case presents consistent findings about the hidden politics of evaluation related to the control exercised by the Federal Government and the embryonic market of evaluation. The nature of the programme, in contrast to SICALIDAD, suggests a more active political use of evaluation findings, but still at a modest level. This section has presented the core findings of the research in relation to the hidden politics of evaluation. The main argument that arises from it is that the differences found in the institutional framework for evaluation allow explaining different manifestations of politics in this process. Even though the rules established in both countries are driven by a distinct logic, the existing discretionary spaces allow political interference. The next section develops in more detail this comparison.

[1123] GESOC represents one of the most active NGOs interested in evaluation. They systematically use evaluation reports to perform their own analysis. For example, GESOC uses evaluation outputs to build an *Index of Performance of Public Programmes* with the purpose of communicating this information to the general public. See: GESOC, A.C. 2014. "Índice de Desempeño de Programas Públicos Federales 2014 (Performance Index of Federal Programmes 2014)." Mexico City.

[1124] Tépach, R. 2014. "Federal Budget Project for the Health Sector 2014-2015." Mexico City: Chamber of Deputies. P. 24.

6.3 Comparative Analysis

The understanding of the political context that underpins evaluation in the UK and Mexico provides a series of elements for the analysis of the way in which the institutional framework relates to this phenomenon. The formal rules set for evaluation in both countries have shown that important differences arise, not only in the definition of mechanisms to shape the behaviour of stakeholders involved in this activity, but also in the way evaluation is conceptualised as a tool of governance.

The hidden politics of evaluation represents a useful means for grasping the relationships established among stakeholders and the balance of power inherent to this interaction. The findings of the four case studies point out those similarities and differences in the institutional framework that affect the nexus between evaluation and politics. This section discusses with more detail the key findings of the empirical evidence obtained and its connection with the theoretical framework (see Chapter 2).

6.3.1 Implications for the Institutional Framework

It seems necessary to look more carefully at the differences found in the case studies. For this purpose, this section discusses five elements of the institutional framework that are affected by the political context in which policies, programmes, and organisations are evaluated or audited in both countries (Table 6.1).

Table 6.1 The Connection between the Institutional Framework and the Hidden Politics of Evaluation

Case	Distinctive Features of the Institutional Framework	Elements Affected by Politics
United Kingdom	No formal integrated policy, general guidelines for evaluation	Lack of enforcement mechanisms of evaluation generates opacity about the current state of this activity in central government and about the quality of the studies and the cost-benefit relationship.
	Wide space of discretion for departments to make decisions about evaluations	Key decisions about evaluation depend on the agenda of departments, focusing on those areas that are considered a priority
	Evidence-based policymaking approach supported by specialised research bodies	The inclusion of multiple stakeholders can be affected by their interests, in a positive way providing more resources for the analysis, but on the negative side it can also bias an evaluation
	Strong faculties given to the NAO to determine VFM of policies, programmes and organisations	The NAO can play an important political role in the discussion of key issues about the performance of government departments.
	Mechanisms of follow-up of the NAO's work by Select Committees	The support of Select Committees provide important leverage to the work of the NAO and to promote the follow-up of recommendations
Mexico	Integrated and enforced evaluation policy	The Executive holds control of the process key elements of the process that might interfere with independence and credibility
	Faculties given to public bodies for coordinating and overviewing the evaluation policy	It generates political struggles that affect the level of influence that evaluation can have in policymaking
	Use of standardised frameworks for evaluation	Limited space for evaluators to act outside the frameworks imposed centrally
	Formal follow-up mechanism for use of recommendations and dissemination of findings derived from evaluation	Rigid frameworks generate incentives for programmes to comply the administrative obligations imposed but with little effect in the overall policymaking process
	Limited role of the SAO for the enforcement of recommendations	Ineffective counterpart to the Executive, limited capacity to make government accountable

The first element regards the existence of an evaluation policy understood as 'high level rules embedded in the legislation used to guide the practice of evaluation.'[1125] The establishment of formal rules for evaluation give more rationality to this process by providing a structured identification of those core elements of this activity, e.g. methodologies, time frameworks, and resources.

[1125] Mark, M. M., L. J. Cooksy, and W. M. K. Trochim. 2009. "Evaluation policy: An introduction and overview." *New Directions for Evaluation* 2009(123):3-11. P. 4.

The cases of the UK and Mexico suggest that evaluation can be a product of different conceptualisations.

While the UK has a long tradition in terms of evaluation and auditing,[1126] in Mexico the adoption of a formal evaluation process is quite recent and is derived from the establishment of regulation in this matter during the first years of the past decade.[1127] In Mexico it is observably a more explicit effort for the institutionalisation of evaluation. The Federal Government has enforced, through regulation, the systematic evaluation of social development programmes, the adoption of a follow-up process for the recommendations, and the dissemination of evaluation reports.[1128] In contrast, the UK has adopted a different approach. There is no integrated evaluation policy that explicitly structures this activity. Key aspects like the selection of programmes to evaluate, the methodologies to apply, and the timescales for studies are left to the discretion of government departments.

The enforcement component of evaluation represents an obvious difference between countries. This can help explaining the different configurations of power that are built around evaluation and the institutional arrangements that can be established. As Jacob et al. recognise, 'Evaluation can follow various designs, is embedded in different forms of institutionalization, and has widely varying usages within different sectors and on different levels.'[1129] Nonetheless, the differences related to the enforcement of evaluation do play an important role in terms of the hidden politics of evaluation. Rules for evaluation determine the amount of power allocated to each stakeholder. In the Mexican case, for

---

[1126] See Lonsdale, J. 2008. "Balancing Independence and Responsiveness: A Practitioner Perspective on the Relationships Shaping Performance Audit." *Evaluation* 14(2):227-48, Morris, S. 2005. "Evaluating Public Policy in the UK: History, Politics & Practice.", Pollitt, C. 1993. "Occasional Excursions: A Brief History of Policy Evaluation in the UK." *Parliamentary Affairs* 46(3):353-62, Talbot, C. 2010a. "Performance in Government. The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom." Washington D.C., US: The World Bank.

[1127] Castro, M. F., G. Lopez-Acevedo, G. Beker Busjeet, and X. Fernandez-Ordonez. 2009. "Mexico's M&E System: Scaling Up from the Sectoral to the National Level." Washington, D.C., US: The World Bank.

[1128] "General guidelines for the evaluation of the federal programmes of the Federal Public Administration." Mexico, "Social Development General Law." Mexico.

[1129] Jacob, S., S. Speer, and J. E. Furubo. 2015. "The institutionalization of evaluation matters: Updating the International Atlas of Evaluation 10 years later." *Evaluation* 21(1):6-31. P. 7.

instance, regulation concentrates the control of the evaluation policy in the Federal Government, specifically in the coordination agencies. In the case of the UK, most of the control remains within departments, while HM Treasury has a limited participation in the process.[1130]

This relates to the existence of a specific body in charge of coordinating and implementing the evaluation policy. The Mexican case is an interesting example, because three different bodies participate in the implementation and overview of the evaluation policy (see Chapter 5, Section 5.1.2). Here, it is worth highlighting CONEVAL's role as the organisation in charge of the evaluation of social policy. Its institutional design as 'a Federal Public Administration decentralized public organization, with autonomy and technical capability to generate objective information on the social policy situation and poverty measurement in Mexico'[1131] has been recognised as an effective way to give independence and credibility to evaluation.[1132]

Despite CONEVAL sitting within the Federal Government structure, the risk of potential political capture is limited by two factors: 'Operating costs (though not the evaluations) financed through a direct budget line in the National Budget; and second, it is governed by an executive board of six independent academics.'[1133] Most of CONEVAL's work presents a positive balance, but claims for more independence have been recurrent in the political debate. Reforms made to the Mexican Constitution in 2014[1134] modified key elements of its governance[1135] and have raised concerns about the risk of political capture:

> The reform states that it is now [...] the Chamber of Deputies, the

---

[1130] This is only for policy evaluation, in terms of performance monitoring the HM Treasury has a much more active role.

[1131] National Council for the Evaluation of the Social Development Policy. 2012a. "About us." Mexico City.

[1132] Sturges, K. M. 2014. "External Evaluation as Contract Work: The Production of Evaluator Identity." *American Journal of Evaluation* 35(3):346-63.

[1133] Gaarder, M. M., and B. Briceño. 2010. "Institutionalisation of Government Evaluation: Balancing Trade-Offs." International Initiative for Impact Evaluation (3ie). P. 9.

[1134] Article 26.

[1135] In the past, Board members were appointed by a commission composed by social development ministers from the states of the Mexican Republic, the Executive and the Legislative; under new regulation, the Board will be appointed by the Chamber of Deputies. Rubí, M. 2014. "Coneval, sin consejeros y sin reglas para elegirlos ("Coneval, no counsellors and no rules for electing them")." in *El Economista*. Mexico City.

institution that will select the academics and this could party-politicize the election. If this happens and the election of counsellors is made based on party quota, the credibility of CONEVAL would be seriously damaged.[1136]

CONEVAL is considered a distinctive feature of Mexico's evaluation system as 'it was the realization of an effort to institutionalize social policy and to give a rather technical than political approach.'[1137] Consequently, potential threats to its independence and credibility are not a minor issue. As its current Secretary states: 'The worst outcome of this political reform would be a totally party-politicised election of counsellors.'[1138] This shows that the establishment of formal rules does not guarantee in practice the elimination of discretionary spaces in the process nor the potential political intromission, as Contandriopoulos recognises:

> The institutionalization of evaluation [...] is a prime component of rationalization policies in countries worldwide. It aims to enhance the performance of interventions by the public sector [...]It is not enough that the demand for efficiency be perceived by public opinion as sufficiently legitimate to have evaluation institutionalized; rather, the various actors must incorporate this new social standard into their ways of conceiving their own responsibilities and those of others.[1139]

In addition, in Mexico the power that the Executive has over key elements of the evaluation process, i.e. scope, periodicity, and methodology, influences greatly the outputs of this activity. This generates an unbalanced distribution of power even within the Executive itself. Coordination agencies hold more power than programme operators or evaluation units within ministries. Regardless of the lack of enforcement, the UK case shows that evaluation has been systematically commissioned or produced internally. However, the extent to which this has been made varies across central government. Recently, the NAO reviewed '6,000 analytical outputs published on departmental websites between 2006 and 2012.'[1140] During 2014, for example, the DoH published 27 independent

---

[1136]National Council for the Evaluation of the Social Development Policy. 2014c. "Factual and legal autonomy: The challenges imposed by the political reform to CONEVAL." Mexico City. P. 4.
[1137] Jaime, E. 2014. "Goodbye CONEVAL?". Mexico City.
[1138] Martínez, N., and A. Rivera. 2015. "Less autonomy for CONEVAL." in *El Universal*. Mexico City.
[1139] Contandriopoulos, A. P. 1999. "Is the institutionalization of evaluation sufficient to guarantee its practice?" *Cadernos de Saúde Pública* 15:253-56. P. 253-254.
[1140] National Audit Office. 2013a. "Evaluation in government." edited by National Audit Office. London, UK. P. 7.

reports about a range of topics such as NHS staff engagement and care for vulnerable groups.[1141] In Mexico, during the same year the Ministry of Health did not publish any reports—neither from compulsory evaluation nor *ad hoc* studies. This can be explained by two reasons: either there were no evaluations carried out during that year or the Ministry of Health failed to meet the obligation of making reports available.[1142]

The existence of a structured evaluation policy contributes to a better control of the process and expected outputs, but at least in the UK's health policy area its absence does not appear to affect the use of evaluation research as it is possible to identify several products commissioned or undertaken by the DoH or by other bodies for the generation of evidence, as well as for the development of guidelines for evaluation (capacity building).[1143] However, it has affected the integration of a system of information about the nature and cost of evaluations, necessary to transparent the use of resources devoted to this activity.[1144]

The institutional framework also determines the extent to which external stakeholders, i.e. NGOs, interest groups, can participate in the process of evaluation. The excessive centralisation of the Mexican evaluation policy leaves a very narrow space for stakeholders outside the governmental sphere to participate. Moreover, the limited market of evaluators constrains the range of options for including an external view. This generates a dynamic in which evaluators become too familiar with the functioning of the evaluation process.

The UK case presents a different situation. Here, there is a broader market of evaluators, specialised research institutions, e.g. NICE, and a range of NGOs and charities actively engaged in the production and discussion of evidence about policy. This gives evaluation more leverage, as more stakeholders are keen to

---

[1141] Department of Health. 2014c. "Publications: Independent Reports." London, UK.

[1142] In the cases of some programmes operated by this Ministry some monitoring record cards were published (including SICALIDAD and Health Caravans); however, this document synthetises information from internal records and does not constitute an independent product of evaluation. National Council for the Evaluation of the Social Development Policy. 2014d. "Monitoring Record Card."

[1143] Interesting examples are the guidelines for evaluation developed by the Nuffield Trust. See: Davies, A., C. Ariti, T. Georghiou, and M. Bardsley. 2015. "Evaluation of complex health and care interventions using retrospective matched control methods." Nuffield Trust.

[1144] National Audit Office. 2013a. "Evaluation in government." edited by National Audit Office. London, UK.

participate. This is also observable in the NAO studies and in the Select Committee inquiries. The existence of formal mechanisms to participate in these processes allows access to stakeholders.

In relation to auditing, both countries have institutionalised systems based on an external body linked to the Legislative with clear rules that determine the characteristics and scope of their work. The role of the Legislative in terms of the follow-up and overview is also clearly established.

The focus of the analysis of both audit bodies is mostly financial and directed to the identification of VFM. The most recent SAO report, for instance, 'identified risks that affect the efficacy and results of public policies and programmes.'[1145] The NAO also establishes the aim of its reports as 'to enhance the degree of confidence of intended users in the financial statements.'[1146] This represents a point of convergence between both countries. Nevertheless, the most notorious difference relies on the political strength of the NAO derived from its relationship with Parliament, particularly with the PAC. These organisations 'are highly influential bodies with government and within wider society because of their high media profile.'[1147] Reports generated through the audit system are an important part of the public debate, as was observed in the UK case studies. The political salience of the PAC gives important leverage to the NAO's work within the public landscape; not only by influencing the audit agenda, but also by highlighting those key findings that are worth a deeper discussion.

In contrast, the political support given to the SAO by the Chamber of Deputies is much more limited. Most of it remains at the discursive level and it is strongly linked to the political agenda of deputies. If an issue of their agenda can be supported by an SAO finding they would tend to consider it. However, most of the time the attention given to the work of the SAO is constrained by the momentum generated by the publication of its annual report. The institutional design of the SAO plays an important role here. The provision of more power

[1145] Supreme Audit Office. 2014b. "General Public Account Report 2013." Mexico City. P. 15.
[1146] National Audit Office. 2015d. "Whole of Government Accounts 2013-14." London, UK.
[1147] Dunleavy, P., C. Gilson, S. Bastow, and J. Tinkler. 2009. "The National Audit Office, the Public Accounts Committee and the risk landscape in UK public policy." London, UK: The Risk and Regulation Advisory Council. P. 5.

and enforcement mechanisms to this body has been an issue of discussion in the public agenda for several years. So far, this has not transcended even though it directly affects the impact of its work. Findings produce very little effect on policymaking because the faculties given to the SAO do not comprise the imposition of sanctions. The institutional design of this body appears to be one of the main reasons for this lack of impact, along with its vulnerability to political influence. According to Ackerman:

> The SAO faces obstacles to perform its control tasks. Among these, it is worth mentioning those political, derived from the pressures of different parliamentary groups; and the economic reflected in an insufficient budget to perform its duties.[1148]

The lack of resources and political support to the SAO has limited its capacity to make the Executive accountable. In the UK, as mentioned, the political strength of the NAO gives recommendations more notoriety, not constrained by the political interest in a particular issue.

The last element to discuss about the institutional framework concerns the extent to which this promotes independence and credibility. Despite the fact that evaluation in the UK is not a structured process and its performance relies strongly on the discretion of departments, it was observed neither in the secondary analysis nor in the interviews performed; no questionings of the independence and credibility of evaluations. In the case of auditing, it was even more notorious that the reputation of the NAO and the PAC is quite strong. The participation of a wider range of stakeholders in these processes appears to have a positive effect in terms of the perception about these values. In the Mexican case, independence is strongly linked to the idea of external evaluation. The evaluation policy is underpinned by the idea that studies are performed by an external expert with an unbiased perspective and that this will generate objective and credible information. This is a core element of the government's discourse. However, the independence of evaluators is constrained by the control exercised by the Federal Government at different stages of the process.

[1148] Ackerman, J. M. 2007. "Hacia un sistema nacional de tranparencia, fiscalización y rendición de cuentas." in *Finanzas Públicas para el Desarrollo*, edited by J. L. Calva. Mexico City: Miguel Ángel Porrúa. P. 256.

The question that arises here is: *To what extent can evaluators be independent if their work is constrained by the decisions made by commissioners?*

In this context, the role of CONEVAL is again relevant for addressing this point. This body has no statutory power, but its institutional credibility comes from its governance. As one of CONEVAL's counsellors (MX-14) explains:

> CONEVAL […] has a lot of impact because of its institutional credibility, it is perceived as a technical qualified institution with academic independence, it has a body of researchers that are not civil servants who have majority in the Board […] the institutional design and the way it has operated is what has given CONEVAL its power.

CONEVAL can be seen as an example of a 'soft power' organisation. It has no mechanism to enforce the use of recommendations but it is a respected institution in the public landscape. It is perceived as a highly competent technical body in the evaluation field.[1149] This becomes particularly relevant at present because of the changes that this organisation will face in the next months. Changes to its institutional design can have a negative effect on its credibility if it is perceived that it has become party-politicised.

These differences between cases show that the configuration of the institutional design does not necessarily isolate evaluation from political influence. Instead, it suggests that this influence can be manifested in different ways according to the rules established for the interaction of stakeholders. For instance, the control held by the Executive over the Mexican evaluation system gives more structure and rationality to the process, but it also interferes with its independence. The value given to evaluation as a mechanism to legitimate public policy gives independence and credibility a more political weight than in other contexts in which evaluation is seen more as a mechanism to support policymaking. This is also affected by the weakness of the SAO, which does not allow establishing an effective checks and balances system. The case of the UK presents an evaluation

---

[1149] According to a recent study, 'there is evidence that CONEVAL has achieved an important position as a body that generates information that is considered a fact when it is disseminated.' Instituto de Innovación Educativa. 2013. "Evaluación Externa de los Resultados del Consejo Nacional de Evaluación de la Política de Desarrollo Social." Mexico City: Instituto Tecnológico y de Estudios Superiores de Monterrey. P. 169.

system more based on the discretion of departments, but at the same time embedded in a context in which external stakeholders play a more active role. Moreover, evaluation is not understood as an accountability mechanism. Instead, it is conceived as a tool for the generation of evidence to support decision-making. The work of NAO is the core element of the accountability system that heavily relies on the political legitimacy given by the PAC. The next section addresses in more detail the relationship between the political nature of evaluation and the institutional framework.

6.3.2 The Hidden Politics of Evaluation: Key Elements of Analysis

The analysis presented suggests that the particular conditions for the performance of evaluation impact on how this activity is undertaken and how it is affected by politics. In this context, it is necessary to look with more detail at particular variables that can explain the hypothesis proposed in this thesis.

The differences observed in both countries related to the institutional framework appear to have an important influence in the way in which stakeholders engage with this activity. A clear example is the perception of stakeholders about the role of evaluation in policymaking. While in Mexico evaluation is seen as a mechanism for legitimising policy, the UK case understands it as an instrument for internal decision-making. Despite the dissimilarities in the purpose, logic, and process of evaluation, in both contexts evaluation does not seem to have transcended its managerial role to feed into the political debate about policy, which, on the contrary, has occurred in the UK with the work of the NAO and the scrutiny made by Select Committees.

The relationship between the institutional framework and the access that it concedes to external stakeholders is a key element to highlight. In Mexico this access is limited and it has produced a biased judgement based on the variables imposed centrally. In contrast, the processes in the UK appear to be more inclusive, and commissioners seem to understand more clearly the advantages of incorporating other views, both for making studies more robust as well as more politically legitimate. This can be explained by the level of discretion allowed by the institutional framework and by the use given to these spaces by

stakeholders. Some of the key variables identified in the analysis are presented in Table 6.2.

Table 6.2 Key Elements of Analysis of the Hidden Politics of Evaluation

Key Elements of Analysis	UK Case	Mexican Case
Legal framework for the performance of evaluations/audits	Evaluation is seen as a mechanism of organisational learning and there is a strong audit system for accountability purposes which is politically supported by the House of Commons.	Evaluation is underpinned by a rigid institutional framework, strongly centralised and that promotes standardised mechanisms of assessment. The audit system is weak, it lacks mechanism of enforcement and there is a total disconnection between evaluation and auditing.
Public sector actors involved in evaluation/audit	In terms of evaluation, the HM Treasury establishes some guidelines, but most of the process is carried out by government departments on their own. In relation to audit, the NAO and Select Committees of the House of Common are the major actors involved.	The evaluation policy is undertaken by three federal government agencies that guide most of the process and some coordination problems have arisen. Evaluation units within government ministries have a more operational function. The SAO is the main body that executes the audit process.
Participation of non-governmental actors	The process of evaluation/audit implies the participation of various actors; mostly limited to evidence provision.	The centralization of the evaluation policy limits the access to non-governmental actors and perpetuates the existence of a small evaluation community.
Nature and scope of evaluation studies	Evaluation studies cover a wide range of topics, mostly related to the improvement of services and the piloting of new approaches to policy. Regarding audit, this activity relies on the VFM approach which aim is to assess the effectiveness and efficiency of public policy.	Most of the evaluation studies undertaken are based on the standardised frameworks promoted by the Federal Government, there is little space for innovation and the approaches to evaluation are mostly related to the implementation, cost-effectiveness and identification of results. The audit system is guided by the values of efficacy and effectiveness and its main target is the identification of deviations from the norm.
Use given to evaluation/audit outputs	Departments are expected to respond to the recommendations made by the NAO or the Select Committees, but there are not obliged to follow them.	There is a formal mechanism for the follow-up of recommendations derived from evaluation; however, it lacks the strength to promote substantial changes in the policy process. Regarding auditing, the SAO has very small capacity to enforce the use of recommendations; most of them are followed but are not substantial.

Table. 6.2 *Continued*

Key Elements of Analysis	UK Case	Mexican Case
Market of evaluators	There is a vast range of researches from different institutions (public/private) that compete for the tenders established for the performance of evaluations, but there are not established mechanisms to assess quality.	It was observed a very small market of evaluators, most of them from public universities and research centres and there are no elements to discriminate in terms of quality.
Perception of objectiveness and independence of assessment	There is no questioning about the independence and objectivity of the assessment, the institutional strength of the NAO legitimises this process. In terms of evaluation, it appears that independence is not a critical issue.	The Federal Government is emphatic on the idea of external evaluation with the aim of preserving the independence and objectivity of the assessment; however, the control over the evaluation process leaves little space for evaluators to make more independent judgements. The role of CONEVAL and its credibility has helped the process of adopting an evaluation model within the Federal Government.
Openness of the evaluation process	The audit process carried out by the NAO and the follow-up enquiries done by Select Committees establish mechanisms for the participation of different actors. There is a vast community of NGOs involved in the inquiries and individuals can also participate.	There are few mechanisms of access for non-governmental actors, mostly promoted by evaluators but supervised by public agencies; there is almost an inexistent participation of the general public.
Linkage to decision-making processes	According to the evidence collected from the field-work, it does not appear that evaluation is an important input in the decision-making; the outputs of auditing seem to have more impact in this area.	Even though there is a formal mechanism for the follow-up recommendations, it is not possible to determine to what extent evaluation is an input in the decision-making process, some minor changes have been documented but other substantial modifications have not been identified.
Effect on public opinion (media and politicisation)	In terms of politicisation, the use of evaluation for this purpose is minimal. In the case of the CQC, the seriousness of its initial failures in events like Winterbourne View provided more notoriety to the scrutiny made to this organisation.	The notoriety reached by some evaluations in the public opinion is related to the relevance of the programme assessed. Only Health Caravans reported some events in which the press focused on some negative features of the programme.

These elements are useful for supporting the hypothesis of this thesis and for answering the secondary questions established (see Chapter 1). The dynamic that is created as a consequence of policy evaluation reflects the different configurations of power among stakeholders, considering the restrictions imposed by formal and informal rules. Taking this into account, the level of maturity that is reached by the evaluation policies in both countries exceeds the formality given to evaluation through the establishment of rules, procedures, and institutions for this purpose.

The existence of institutional channels for the participation of stakeholders and a more solid checks and balances system has generated more favourable conditions for the blooming of evaluation and auditing in the UK, regardless of the lack of an integral and central evaluation policy. In the Mexican case, it is observable that important efforts for institutionalising evaluation and for generating capabilities in this matter have been made, both by the Federal Government as well by others, i.e. international organisations. Despite this, the way in which this process is controlled centrally has limited the utility of evaluation to a programmatic level that does not reflect completely its contribution as a tool of governance.

In both cases, the analysis of these countries uncovers the relationship between evaluation and its political context and those key elements of evaluation in which political influence is observable. The final section of this chapter elucidates the connection between the empirical findings and the theory.

6.3.3 Empirical Evidence about the Hidden Politics of Evaluation: The Connection with the Theory

The relevance of studying the relation between evaluation and politics is contributing to the existent knowledge in this field. The empirical evidence obtained allows looking at the different ways in which stakeholders establish political linkages in evaluation. Considering the intangibility of politics and the multiple factors involved, this research has shed new light on some key variables of this process in which politics intervene, but also on the way in which institutional frameworks lead to different configurations of power. In this

context, it seems necessary to identify the linkage between the empirical evidence and the theoretical knowledge about the hidden politics of evaluation to make more explicit the contribution of this thesis.

According to the theoretical framework (see Chapter 2), the core ideas about the political nature of evaluation are essentially related to the interaction between stakeholders and the pursuit of their interests. One of the most distinctive features is reflecting upon how stakeholders (particularly evaluators) can cope with the political context in which this activity takes place. The approach of this research has allowed broadening the scope for analysing this phenomenon in a more holistic way. The discussions presented in this chapter are not constrained to the role of evaluators and their interaction with other actors, but to the understanding of the effect of the institutional framework.

Beyond the negative connotation given to politics, this research has, from its inception, acknowledged the idea that evaluation and politics cannot nor should be separated. With this in mind, the theoretical and empirical analysis has demonstrated that the emphasis should not be on developing new ideas about how to cope with politics in evaluation (considering it as a menace to values like independence and credibility), but to find new ways for grasping this relationship and its meaning for the role of evaluation in governance.

The empirical evidence is consistent with the theory by showing that the control of an evaluation policy relies largely on the capacity of stakeholders to use their resources towards an end and the different degrees of influence they can exercise. Consequently, formal and informal rules for shaping the behaviour of stakeholders will affect their interaction. The institutionalisation of evaluation demands transcending the establishment of procedures to a more clear comprehension of the roles that stakeholders will have, and the potential gains/losses of power that might take place.

Considering governance conditions, stakeholders could see more clearly the potential benefits of evaluation as a tool for strengthening their position in the

political debate about the effectiveness of public policy. Here, the empirical evidence obtained is relevant because it shows that even though institutional arrangements differ, the use of evaluation as a governance tool is modest in both cases.

The Mexican case, for example, suggests that the isolation of the evaluation process and its excessive centralisation keeps the political struggle within the public sector. The most evident struggle can be found in the resistance of programme operators to be evaluated and to disclose information for evaluation. Disagreements produced are mostly resolved by the statutory power of the Federal Government to enforce cooperation. Consequently, the relevance of evaluation as a political tool is limited and stakeholders mostly perceive it as a managerial tool.

The values promoted through evaluation are another empirical element worth highlighting. Although these values are not always explicit, there is a clear bias towards the identification of the effectiveness of policy in both countries, based on the relationship between inputs and outputs. The values promoted by each stakeholder, according to the theory, can be diverse and at some point contradictory. It was observed in the case studies that efforts of audit/evaluation aim to obtain evidence about the cost-effectiveness of policy, even when other actors, e.g. charities and NGOs, consider that other values should be assessed. The rationality of evaluation in terms of politics concerns the capacity of stakeholders to pursue the predomination of their interests. In both cases analysed, the strongest actor was the one with enough political power to control the process. In the UK case, the role of the NAO is associated with accountability. Its work in the identification of the VFM of government programmes is acknowledged as relevant evidence for discussion in the political arena. In the Mexican case, the control of the evaluation process is concentrated in the Federal Government. In contrast to what happens in the UK with the NAO, the control exercised by the Mexican government is more perceived in the rigidity of the norms and procedures established for evaluation, rather than in its political strength and legitimacy, and also this control is affected by power disputes among coordination agencies (see Chapter 5, Section 5.1.1).

Regarding the rationale of evaluation, the evidence suggests that there is a different conceptualisation of the role of evaluation. In the UK it is seen as a mechanism to strengthen the decision-making (organisational learning), for example, through the performance of pilot schemes for a new policy. In Mexico, evaluation is more related to the identification of goals achieved and this is strongly associated with an idea of legitimisation of public policy. The Federal Government aims to make explicit the external nature of evaluation to underline the independence and objectiveness of the process, even though the assessment is strictly controlled internally.

As the theory suggests, there is no possible way to abstract evaluation from its political context. Evaluators and practitioners must operate in an environment in which stakeholders are constantly struggling for their interests. The findings from the fieldwork show that the role of evaluators becomes especially important in the selection and analysis of evidence based on their capacity to detect when stakeholders are willing to bias the process. The case of the CQC illustrates this point. The controversial events in which this regulator was involved during its first years provoked a more intense scrutiny of its performance and more public attention. The inquiries made by Select Committees showed that the prioritisation of evidence to be discussed by MPs is an area subject to the discretion of the actor that selects it.

In addition, the access to participation channels in evaluation processes is also a political matter. There can be different levels of inclusion of stakeholders in evaluation. While most of the evaluation processes imply the participation of different actors, this does not mean that they have an active role. Consequently, not all relevant views are considered. The four case studies showed that there is still a long way to go in terms of the democratisation of evaluation. Although there is a more open access in the UK for external stakeholders than in Mexico, their participation is still limited. Mostly, it is reduced to the provision of evidence with little room for debating.

An interesting finding related to the theory is the interaction between stakeholders. In the UK case, for example, the relationship between the NAO and

departmental officials reinforces the idea about how actors can take advantage of discretionary spaces in the process. VFM studies imply the interaction of people in the audited body and staff from the NAO and this cannot, nor should be completely normed by the institutional framework, and it is natural to find some discretionary spaces. A more concrete example is the idea about how VFM studies are built around 'agreed evidence'. This does not mean *de facto* that there is an express intention to manipulate the assessment, but stakeholders can affect the outputs. In this context—data-collection—the timing for evaluation and the methodology represent only some examples of decisions that can be negotiated (generally informally).

The political component of evaluation does not refer to the obscure and negative image associated to partisan politics, it concerns the interaction between stakeholders that takes place during evaluation, constrained by the resources available and the conditions set by the institutional framework. This connection between theoretical and empirical evidence demonstrates that it is not possible to abstract evaluation from its political context. However, the core element of discussion is the way in which politics is managed. One of the most relevant findings produced by this research is the influence exercised by the institutional framework in this politics/evaluation relationship. Different contexts can generate multiple configurations of power, the level of access can produce different outputs, and the dynamics in which evaluators and commissioners relate can also affect the way in which evaluations are conducted.

The next chapter synthetises and presents the most relevant conclusions of this research, as well as some recommendations for coping with the political nature of evaluation.

Chapter 7. Conclusions and Policy Recommendations

This research has elucidated the way in which politics and evaluation interact. The careful study of two different contexts produced an important range of findings worth highlighting in this final chapter. In both countries (UK and Mexico) evaluation has become an instrument with multiple applications. In all case studies, the production and use of evidence for judging different dimensions of policy has been a common element. However, the characteristics of this evidence and its level of utility are aspects that differ between countries. Yet, the core issue of this thesis is not the instrumental use of evaluation in government, but its political nature, i.e. the way in which it manifests in the public sphere and its consequences for the governance model. Uncovering the hidden politics of evaluation through the case studies has enlightened the path for understanding this phenomenon. The negative connotation given to politics (in the partisan sense) contrasts with its potential benefits as an instrument for fostering changes and reaching (negotiating) collective goals.

As discussed, the rules and procedures established for evaluation and auditing shape the behaviour of the stakeholders involved in these processes. This becomes particularly relevant in terms of the dissemination of responsibilities and obligations in this matter among different actors. These conditions affect the different configurations of power that stakeholders establish in the pursuit of their interests. The potential impact that evaluation outputs can have on their interests might help explaining how they engage with this activity. Nonetheless, the institutional architecture of evaluation is unable to eliminate completely the discretion of stakeholders in the process. In this sense, as a political instrument, evaluation relies on the capacity of stakeholders to take advantage of power spaces found in the process for pursuing their agendas. This last chapter reflects back upon the findings obtained to make explicit the theoretical and empirical contributions of this thesis. For this purpose, it is divided into three sections. The first one presents the main conclusions derived from the analysis in both countries, emphasising the most notorious differences found. The second section explains the theoretical contribution of this research: *To what extent has*

this thesis expanded the frontiers of the existing knowledge about the hidden politics of evaluation? Finally, a series of recommendations about how the political nature of evaluation can be addressed in the theoretical and practical domains are presented. These reflections aim to explain to what extent the use of evaluation has made the state smarter, in the sense of making the most of this managerial and political instrument—in other words, adopting evaluation mechanisms, taking into account the challenges associated with the interaction among the participants of this activity.

7.1 General Conclusions of the Research

At a broader level, the study of the hidden politics of evaluation shows that the evidence obtained in both countries presents some common characteristics. Considering that the socio-economic and political characteristics of each country have promoted a different approach to policy evaluation (see Chapter 3, Section 3.2.3). There are some major elements of the discussion of the political nature of evaluation that transcend these differences. This section presents some mainstream arguments that concern fundamentally the role of the institutionalisation of evaluation as a variable that can predict (partially) the political behaviour of the stakeholders in evaluation.

The first argument to highlight is that despite the differences in the institutional design for evaluation, the formalisation of this activity in both contexts addresses key issues in this matter: *Who will perform evaluations? Which instruments will be applied? How will funding be allocated?* The degree of this formalisation differs in terms of the rigidness or flexibility of the norms (or guidelines) for this purpose. From an overall perspective, the perks of giving evaluation more structure are mostly constrained to the managerial aspect. The political nature of evaluation seems to be hidden behind the technical and rational dimensions associated with this activity.

Another interesting element in common refers to the disconnection between monitoring, evaluation, and audit systems. There are systems in place operating simultaneously; yet its complementarity (and perhaps overlapping) has not been analysed in depth. For this reason, it is complex to determine the extent to

which these activities are contributing individually and collectively to the improvement of policymaking.[1150] This matters in terms of how evaluation feeds into accountability systems. Its contribution needs to be explicit, for example by stating how policymakers will use evaluation outputs to inform and justify their decisions. This resonates with the discussion about evaluation contributing to the reduction of the democratic deficit (see Chapter 2, Section 2.1.1). The promotion of evaluation approaches that promote transparency and participation of other stakeholders can help narrowing the gap between citizens and government. Moreover, it highlights the link between evaluation and auditing. The description of the institutional framework for these activities in both countries suggests that the differences between them are not only procedural but also epistemological. Auditing has a political value related to its contribution to accountability, while evaluation is perceived mostly as a managerial tool. These different conceptualisations are reflected on; the extent to which the institutional framework promotes the use of outputs derived from this activity, i.e. follow-up inquiries by Select Committees.

Quality is an issue that stands out in both contexts. However, the effect of this variable differs. The rules and procedures set for evaluation do not appear to determine (at least not explicitly) clear criteria to verify the quality and robustness of evaluation. Despite this, the questionings to credibility seem to be more perceivable in the Mexican case. This leads to reflecting about how power and control in an evaluation system should be allocated. In both scenarios, government has a predominant position in relation to other stakeholders. This imposes important challenges because it limits their participation in key issues, like assessing the quality of evaluation. What is more, in terms of credibility and independence, the dissemination of the control of the evaluation process might increase the trust on it and its utility for stakeholders.

A related concern regards the value given to external evaluation seen as a safeguard for independence and objectivity. The discussion presented about the internal versus external evaluation models shows that independence relies on

---

[1150] In the Mexican case, for example, the fact that three different organisations conduct the evaluation policy has generated disputes among them that affect the conduction of monitoring and evaluation systems.

more variables other than this choice. Nevertheless, the cases analysed show that there needs to be more explicit criteria for the commissioning of a study to justify the rationale of selecting an in-house or an external approach to evaluation.

Lastly, the acknowledgement of evaluation as a risk-reduction mechanism requires reflecting about what this means for the stakeholders involved. From an instrumental perspective, this would refer, for example, to giving policymakers evidence to design and implement policies in a more controlled environment. And yet, the connection to the interests to other stakeholders is not that straightforward. Politicians, for example, do not seem to be completely aware of the potential of evaluation outputs for the development of their political manifestos or for the identification of political priorities. In governance, this reveals that evaluation needs to expand the boundaries of its rationale to incorporate the needs and expectations of actors beyond the governmental sphere. It is certain that the governance model demands recognising the existence and influence of other actors and, in this sense, evaluation can be adopted as a mechanism for helping them gaining new spaces of power in the public sphere, for example by giving civil society a more visible position in the debate about the effectiveness of government policies (democratic evaluation). Also, this could mean increasing the influence of evaluation bodies (not necessarily governmental) as entities devoted to the production and analysis of evidence that can be used for political purposes. This would mean recognising evaluation as a tool of power and influence and reflecting on who should lead an evaluation policy.

7.1.1 Conclusions about the Hidden Politics of Evaluation

The case studies analysed allowed establishing a series of conclusions about how politics and evaluation interact in two different contexts. These findings highlight those keys elements of the research that contribute to generating new knowledge in the evaluation studies field. As stated, the hidden politics of evaluation can be perceived in both contexts as the exercise of power and resources by stakeholders to influence the evaluation process, according to their interests. The political nature of evaluation might not be explicitly

recognised by the institutional framework, but it is easily identifiable in the formal and informal arrangements that take place in the conduction of this activity.

In relation to the control of the process of evaluation, an interesting finding is the existence of a predominant actor with the capacity to shape the process according to his interests. Under the light of the thematic framework proposed in this thesis, this suggests that the control over key elements of the process, e.g. purpose, resources, etc., is what allows this actor to maintain a more favourable power position. While in Mexico this actor is clearly the Federal Government, in the case of the UK, evaluation lacks a predominant actor as responsibilities in this matter rest at departmental level. Nonetheless, this is not the case in auditing, for which the NAO, and in a broader sense, Parliament, are the leading organisations. These differences can be explained at some level, because of the asymmetries in the political leverage given to evaluation and auditing.

Regarding the institutional design for evaluation, it is possible to observe that this has been the result of the political arrangements that finally led to the establishment of rules and procedures. In the Mexican case, for example, the role of CONEVAL stands out as an evaluation agency with which legitimacy is based on its technical expertise and on the 'soft power' that it exercises. Despite this, the political changes that recently occurred in Mexico will eventually promote a new configuration of power and political arrangements that can undermine the legitimacy of this body and that will need to be carefully studied. This contrasts with the UK, where the role of the NAO and the PAC in the public landscape are more intrinsically recognised and trusted as scrutiny bodies.

The political nature of evaluation, contrary to intuition, has little to do with the manipulation of outputs or the use of evidence to pursue particular interests. What this research suggests is that the politicisation of evaluation is mainly manifested in the influence exercised by stakeholders during the process in the control of key stages of evaluation, i.e. data-collection and method-selection. This relates to the struggle for dominating the 'evaluation space' and the outputs of the relationships produced in it. The resulting outcome of this is the

predomination of the interests of the actor that holds more power, i.e. the Federal Government in Mexico. Considering the relevance of the institutional framework, this influence reveals that stakeholders take advantage of those spaces of discretion (not formally regulated) to exercise their power.

Another important conclusion is that neither in Mexico nor in the UK has evaluation reached a level of notoriety to promote its inclusion in the political debate about the effectiveness of government. This relates to the technocratic value given to evaluation which seems to have constrained it mostly to the managerial sphere. As a consequence, its potential as a political tool seems underestimated. The depoliticisation ideas discussed in the theoretical framework (Chapter 2, Section 2.1.1) are useful for underlining the relevance of the rationality label given to evaluation in the aim of detaching it from its political nature. This can be perceived, for example, in the way in which the Mexican Federal Government emphasises the 'external' nature of its evaluation policy.

As a tool of governance, evaluation seems to be still at an immature stage. More specifically, its utility for making more informed decisions, generating better conditions to negotiate with stakeholders, and identifying those responsible for public policy is unclear in the cases analysed. It could be expected that this utility would increase according to the level of influence that stakeholders can exercise over the process. For instance, as a political tool for avoiding blame, the utility of evaluation will depend on the capacity of stakeholders to participate in the process of evaluation without this being perceived as an open attempt of manipulation. What arises from these reflections is that the control over the 'evaluation space' implies cooperation among stakeholders, for the preservation of the legitimacy of the evaluation process. Undermining the credibility of the evaluation system represents a threat for all stakeholders involved, regardless of their individual agendas. The relationship between the 'evaluation space' and Dubnick's concept of 'accountability space' serves the purpose of highlighting the importance of studying how control mechanisms interact with the political context from a more holistic perspective. This means looking in detail at the relationships that emerge among different actors in the design and

implementation of evaluation processes.

These conclusions provide the necessary basis for reflecting back on the research questions (Table 7.1). The hypothesis proposed in this thesis is that the institutional framework designed for an evaluation policy can determine the degree of political influence in evaluation and may affect the consequences for the utility and legitimacy of this activity. The evidence discussed here suggests that the institutional architecture for evaluation defines the configurations of power that are created. This can be observed in the way rules and procedures determine the degree of independence that evaluators have; the prerogative for commissioning evaluations and the mechanisms for the dissemination of findings, among other core elements of the process that affect the outputs and outcomes.

The conclusions of each one of the cases discuss these issues in more detail. However, the point to highlight here is that the political nature of evaluation was observed in both countries. The major difference was the way in which this nature manifests due to the conditions provided by the institutional framework and how this shapes and constrains the behaviour of stakeholders.

Table 7.1 Linkage between Core Questions of the Research and the Conclusions Obtained

	Research Questions	Main Conclusions
Core Question	What is the relationship between politics and evaluation?	It was acknowledged that evaluation is a political activity *per se.* The implications of this were observed through the analysis of the case studies. In different stages of the process of evaluation/audit there were political issues at stake. Most of these issues were related to the control of the process and the predomination of the interests of the actor in this position.
Secondary Questions	Are there efficient institutional frameworks that assure an effective assessment of public policy?	In both countries there are institutional arrangements for the performance of audits and evaluations. Even though important differences arise (e.g. departments' freedom to evaluate), there are formal processes related to the evaluation of public policy. The main conclusion is that in both scenarios there are spaces for stakeholders to act with discretion which is where the politics manifest more clearly (see Chapter 6).
	Which values are promoted by the different stakeholders involved in evaluation? How do these values interact?	The values underpinning evaluation and auditing identified in both countries are related to efficiency, effectiveness and economy. These values are aligned within the formal structures in which these activities take place; however, from the perspective of external stakeholders other values should be put at the core of evaluation and audit processes. In the case of health policy, for instance, equity and access were some of the values underlined by these actors.
	Does the evaluation policy, in both countries, produce in a systematic way relevant and convenient information that can be used for the policymaking process?	In both countries there is a systematic production of evaluation outputs derived from the work of different actors. In the UK cases, for example, it was observed a more plural participation of actors in the production of new knowledge, while in Mexico this is an activity mostly monopolised by the Federal Government.
	When the outputs of evaluation are used, is there any effect on the improvement of policymaking?	In the four case studies it was identified a series of evaluation outputs and some follow-up processes. Specifically in the Cancer Strategy, the use of evidence is a recurrent practice due to the nature of this policy issue. In the rest of the cases, the influence of the information derived from assessment processes is difficult to quantify, some examples were provided but there is no unquestionable way to determine the effect of evaluation into the policymaking process.
	Which is the role of policy evaluation in order to consider it useful for the policymaking process?	Here an important difference arises between the two countries. While in Mexico the Federal Government has given evaluation a primary role as a mechanism of legitimisation, in the UK the use of evaluation research is more perceived as a tool for policy improvement and organisational learning, while the audit system that operates exercises important accountability functions.
	In terms of 'good policy' and 'good politics', which of these issues predominate in the evaluation agenda of both countries?	In both countries evaluation plays an important role in public management; but it has not transcended to political discussions. Consequently, the politics of evaluation are mostly constrained to the use of power and influence in different stages of the process, rather than to the use of evaluation as a systematic input for the construction of political arguments.

7.1.2 Conclusions of the UK Case

Most of the differences in relation to the Mexican case regarded the set of conditions and rules established by the institutional framework. The existence of a parliamentary system—as opposed to the Presidential regime in Mexico—promotes a more plural and inclusive political landscape (checks and balances). In the UK, there is a more equitable allocation that impedes—at least formally—an excessive concentration of power in one actor. The strength of the UK audit system represents another important distinctive variable. Considering this, the way in which the political nature of evaluation is manifested varies importantly in respect of the Mexican case.

The rationale for evaluation in the UK is perceived more as an organisational learning asset than as a means for policy legitimisation. The lack of enforcement mechanisms for the performance of evaluation gives government departments discretion to carry out studies according to their needs and resources. Most of these studies are considered inputs for the internal decision-making and a tool for constantly reflecting upon possible courses of action. This finds echo in the discussion about the need for risk-reduction and control instruments demanded by governance. In this sense, a more informed policy aims to reduce the controversy around government decisions and to generate more political support. In particular, this is the case of the Cancer Strategy, as a policy strongly based on scientific evidence, the effectiveness of which focuses on the values being prioritised by key stakeholders, while the instruments for the implementation of governmental action are less controversial because of its evidence-based nature.

Another interesting conclusion of the UK case concerns the power held by the NAO. This is perhaps one of the most substantial differences between evaluation and auditing that can help explaining the gap existing between these activities in relation to their effect on policymaking. The political support provided by the PAC gives the NAO the strength to effectively scrutinise how government departments are delivering services and to identify the VFM that they are producing. In addition, its institutional design favours the construction of credibility in the NAO and the statutory power to promote changes in the public

sector.

The fact that there is more freedom in the way evaluation is performed in the UK does not mean *de facto* that there is more discretion in this process. Core elements like the commissioning of evaluations are regulated and this removes from commissioners the decision to select evaluators in a discretionary way. Moreover, the guidelines produced by HM Treasury allow government departments to adjust evaluation to their particular needs; this might help incentivising the use of social research as an aid for policymakers. This flexibility might be explained by the fact that in this country the perception about the independence and objectivity of evaluation is not a major issue. As a result, more discretion is not perceived negatively.

The existence of an active research community, charities, NGOs, think tanks, and other actors increases the plurality of policy evaluation. The way in which the institutional framework allows the participation of these actors favours the inclusion of different views about a particular policy issue, and yet, the extent to which they are able to participate is limited. In line with this, the democratisation of evaluation could be seen as way to open this process and incorporate other views. This could even imply a collaborative design of evaluation instruments.

In more specific terms, the analysis of the Cancer Strategy case provided evidence about how the participation of stakeholders can become political. For example, different disease groups will intervene, driven by their particular interests. Although cancer is considered a highly technical policy issue, there are spaces for stakeholders to seek the predomination of their interests, e.g. more funding for a particular type of cancer. The political notoriety of the Cancer Strategy and its inclusion in the political agenda of the government and MPs give more salience to its evaluation findings and the political use given to these, e.g. for the development of political manifestos in electoral times.

The case of the CQC offers a different perspective about the hidden politics of evaluation. The widely known shortcomings in its performance at the beginning

of its operation increased the pressure for the evaluation of this organisation. According to the evidence collected, the scrutiny for the CQC has been more strict and the follow-up more persistent. This reflects that the drivers for evaluation can be highly political. If there is a public negative opinion about an organisation, scrutiny bodies are keen to be more critical, to promote the adoption of changes and to make sure that recommendations are being heeded. The accountability function of evaluation in governance refers precisely to the identification and acknowledgement of the expectations created around an organisation. Considering this, the political relevance of the evaluation of the CQC, after its troubled first years of operation, generated momentum for the adoption of this activity as a core element of its organisational behaviour.

The use of the evidence produced by the NAO, the PAC, and the Health Committee about the CQC shows that an evaluation can become politicised for several reasons. In this case, the failures in addressing key events and the organisational differences between members of the Board led to a more severe scrutiny, more public attention, and the distortion of the main focus of the CQC. Here, evaluation can become a counterpart for the potential blame-avoidance by providing robust evidence about the performance of this regulator.

7.1.3 Conclusions of the Mexican Case

The explicit promotion given to evaluation by the Federal Government in Mexico has led to the development of an evaluation policy with which implementation has generated important outcomes in a relatively short period of time (see Chapter 5). The Mexican case presents interesting findings about the institutionalisation of evaluation in the context of a country with an important legitimacy crisis of public trust.

Despite the fact that the Federal Government has strongly promoted the adoption of an evaluation policy, the institutional framework underpinning it seems weak and still at an emerging stage. The audit system currently operating shares some of the same problems observed in evaluation. The scope of audits is fundamentally procedural. Consequently, the changes and modifications that can be produced through this system are minor. In addition, the lack of

enforcement mechanisms and the political power of the SAO to promote the adoption of more substantial reforms or to sanction the underachievement of goals has limited the impact of this activity. The rigidness of the evaluation and audit systems does not allow other stakeholders to access the 'evaluation space'; they are constrained to a passive role that limits also the utility of evaluation outputs for them.

The role given to evaluation in Mexico is more related to its use as a legitimation tool; not entirely as a mechanism for the adoption of reforms. This can be mostly observed through the emphasis given to the idea of 'external evaluation'. Although the Federal Government strongly controls the evaluation process and limits the tasks of evaluators, it seems very important to this actor to make explicit the fact that external parties are the ones providing the judgement. This is helpful for discussing the pertinence of giving evaluation more political value. Especially in the Mexican context, in which credibility and independence are core issues at stake, giving more political salience to evaluation might generate costs associated with the perception of a biased judgement of policy. The technocratic value given to evaluation seems to fall into conflict with the promotion of a more explicit political approach to evaluation.

As stated, Mexico's excessively centralistic evaluation policy leaves very little room for the participation of non-governmental actors or even actors within the government that are excluded from key decisions, e.g. operators. This manifests the influence of the institutional framework in the distribution of tasks in the evaluation process and the rationale behind it. For instance, removing commissioning decisions from ministries and transferring them to coordination agencies (depoliticisation) aims to establish distance between evaluators and evaluands. However, this distance seems also to affect the utility of evaluation for programmes.

This can be observed in the case of the Health Caravans Programme, with which disagreements about the scope of evaluation between decision-makers and operators have occurred. The distance perceived between those who set the evaluation agenda and the knowledge about the operation of the programme

was reflected in the perception about the lack of utility of evaluation. Programme operators, for example, are perceived as mere executors of a decision taken outside the core sphere of the programme. Health Caravans demand an important amount of information from the field which cannot be obtained through the desk-based evaluations that are constantly performed.

This case is also relevant for discussing the relationship between the political notoriety of the programme and the political use of evaluation outputs. The promotion of the achievements of the programme and the validation of political decisions like the modification (reduction) of budget can be supported by the evidence produced by evaluation.

In relation to SICALIDAD, the ideas around the hidden politics of evaluation are linked to the use and sharing of information. As it constitutes the main input for the evaluation, the control of the information is a way to influence the process. Evidence suggests that in line with the theory, evaluation can be perceived as a potential threat for the interests of stakeholders. Consequently, there is a patrimonial vision of the information in the aim of diminishing the risk of a possible negative output. This, in the context of governance, can be linked to the need for cooperation and the generation of incentives for aligning the behaviour of stakeholders. This cooperation might only be achieved through the flexibility of the evaluation process and the delegation of some of the functions from the Federal Government to other stakeholders.

SICALIDAD also illustrates how discretionary spaces found in the process of evaluation can be used by stakeholders in the pursuit of their interests. This does not necessarily imply an explicit intention to manipulate the process, but it can certainly affect the output of an evaluation, e.g. by modifying the selection of a sample that can lead to different findings. This makes evident that although a rigid process is implemented for the design, commissioning, and supervision of evaluation, rules, and procedures cannot completely predict and control the decisions of stakeholders.

What the Mexican case shows is that the actor (Federal Government) that

controls most of the process can subtly shape it according to its interests, while at the same time it can control the rationale of evaluation for the legitimisation of public action.

7.2 Contributions to the Theory

The theoretical contribution of this thesis to the knowledge about the hidden politics of evaluation is relevant for two main reasons. The first one is that obtaining empirical evidence about politics helps identifying some new areas of study in the field. Secondly, as most of the theoretical production about the political nature of evaluation has focused on the struggle of interests among the stakeholders, it seems necessary to expand these boundaries in relation to how the role of evaluation in governance can be affected by politics. The aim of this thesis has been to push the boundaries of the existent knowledge by looking at how the conditions established by the institutional architecture affect the relationships that take place in the 'evaluation space'.

The ideas about evaluation as a completely objective and apolitical instrument were reconsidered later by the scholars as the evidence showed that it was not possible to abstract evaluation from its political context. These ideas were mostly based on the fact that the objects of evaluation are the result of political decisions, when the people involved in evaluation (practitioners, operators, evaluators, or politicians) have political values to defend, and the consequences of the assessment would eventually lead to a new set of political arrangements, e.g. distribution of resources, focus expansion, or programme termination. The way in which the evaluation is embedded in a political dynamic became an issue of interest for the scholars, to understand how this might affect evaluation.

The potential contributions of this thesis to the theory are addressed in this section, taking into account the theoretical pillars that guided this thesis.

7.2.1 Governance, Policy Analysis, and Performance Management

The new functions acquired and developed by governments around the world are strongly related to the new challenges imposed by the reality. The adaptability of governments to these changes demands the adoption of new

managerial models for policymaking. One of the main findings of this research is the use of evaluation processes as a mechanism to respond to two major demands: the development of more effective policies and for more accountability.

In this context, this thesis has generated empirical evidence about the interaction of the different stakeholders involved in evaluation. Through this evidence, it is possible to understand to what extent governments have been forced to rethink their roles, due to the external pressures. Moreover, evaluation and auditing have also had an important effect within public organisations, i.e. resistance to evaluation, that show that the participation of non-governmental actors is not passive. Instead, they are taking new positions of power that allow them to promote their interests. The main theoretical contribution of this research to the governance and policy analysis can be summarised into two major arguments: the first one is that the transformation of the state goes further in the implementation and adoption of new models, e.g. NPM, in which evaluation has the potential to transcend its role as a managerial tool to become an instrument for the construction of political arguments. The acknowledgement of the politics of evaluation can be seen as a process for developing new institutional arrangements that can broaden the participation of non-governmental actors. This view is based on the argument that the rise of new stakeholders' demands has promoted thinking about their participation, not only in evaluation but also in the whole policymaking process. Likewise, the changes in the conceptualisation of the role of evaluation are now more directed towards the inclusion of multiple views (democratisation). Consequently, a more inclusive approach to evaluation might fit better into the context of governance, not only because of its benefits for the robustness and credibility of evaluation, but also in terms of its capacity, to contribute to the reduction of the democratic deficit.

The other argument is that looking at evaluation as a stage of the policy cycle whose main function is to verify the accomplishments made by a programme, dismisses its potential as a key element for the generation of evidence susceptible to be used at any other stage (agenda, implementation). Therefore,

understanding evaluation merely as a component of the policymaking process limits the possible benefits to be derived from this activity. Its institutionalisation needs to take into account not only its role as an aid for policymakers, but also its potential as a political tool within and outside the governmental sphere. The technocratic nature of evaluation needs to be transformed, in order to build its political legitimacy. As a tool of governance, evaluation demands to adapt its approaches to a new dynamic, which entails a more intensive use of this mechanism. This implies understanding how stakeholders engage in the evaluation process and how they take advantage of the spaces found in the institutional architecture.

In relation to performance management, the findings derived from this research resonate with the theory behind this model, at least in three different ways. First, the interaction of performance-management systems and other control mechanisms, i.e. evaluation and auditing, implies the coexistence of values that are not always aligned. This generates fragmentation that makes difficult the identification of the overall contribution of these instruments to policymaking, and the extent to which these activities are complementary or supplementary.

Secondly, the necessary conditions for the operation of these performance-management systems are not produced automatically. For example, decentralisation of government, through the creation of arm's-length bodies, might be driven by the idea of increasing efficiency as well as promoting specialisation or independence. However, it creates some other distortions that cannot be ignored, such as the difficulties associated with establishing effective mechanisms of control. This becomes even more complex when the institutional framework is ambiguous, as in the case of the coordination agencies in Mexico.

Finally, the evidence generated by this research shows that the simultaneous application of control instruments does not guarantee an effective appraisal of policy. The theoretical issue to discuss here is the extent to which the coexistence of different assessment systems can produce more harm than good. Designed and underpinned by a different logic, it is worth reflecting on their role in the broader context of governance; particularly, in terms of how the

politics embedded in these activities is recognised and addressed. As an input for evaluation, for instance, performance indicators might imply a bias not explicitly identified but that can affect the overall judgement.

7.2.2 Evaluation Studies

Looking at evaluation studies as a broader field of knowledge, the development of this thesis has also discussed different arguments about the theory and practice of evaluation that can guide future studies in this field. On the one hand, this thesis synthesises the most recent discussions about the conceptualisation of evaluation and its utility for policymaking. On the other hand, it has also elucidated the relevance of taking into consideration the different rules and institutions that are created for the performance of this activity in the public sector.

In line with this, a contribution of this thesis to the evaluation studies field can be seen in the identification of the institutional framework as a core element of evaluation practice that affects substantially this process. Moreover, the multiple ways in which formal and informal rules can shape the performance of evaluation can lead to different results. This is especially evident in the arrangements made among stakeholders, for instance in the way they agree to heed recommendations.

Another contribution is the recognition of core elements of the process of evaluation that are affected by the interaction among stakeholders. Evaluation is not a unilateral activity; it involves the participation of stakeholders within and outside the governmental sphere whose interests can determine how this activity is conducted. The practice of evaluation seen from the political perspective demands reconciling the interests involved in this process and identifying which areas need to be formally regulated to guarantee its robustness, credibility, and utility.

7.2.3 Theory about the Hidden Politics of Evaluation

Scholars in the evaluation studies field have made an important contribution acknowledging the political nature of evaluation and the implications for the field.[1151] In particular, these discussions have concerned the way in which politics can affect the work of evaluators and how they can cope with it to preserve the integrity of their work.[1152] Although the challenges faced by evaluators and practitioners are of great importance, this research has aimed to go further by understanding the relationship between evaluation and politics through the lenses of the institutional framework in which this activity takes place.

The main contributions of this thesis can be identified in three main issues. The first one is the development of a thematic framework (see Chapter 2) whose purpose was to provide an analytical tool to grasp a complex and intangible topic like politics. This framework identified five variables (purpose, resources, process, outputs, and outcomes) considered crucial for guiding the research. Through this tool it is possible to establish some basis for a more structured comprehension of how politics and evaluation interact. It is neither exhaustive nor fixed, so it can be improved and adapted according to the needs of new objects of study.

The second contribution is the empirical evidence about the hidden politics of evaluation that allowed the identification of new variables to explore (see Chapter 6, Section 6.3.2) to reach a more specific level of analysis and more data that can feed into new research models.

Finally, the identification of key themes of the hidden politics of evaluation that

---

[1151] For example: Chelimsky, E. 1995. "The Political Environment of Evaluation and What it Means for the Development of the Field: Evaluation for a New Century: A Global Perspective." *American Journal of Evaluation* 16(3):215-25, Palumbo, D. J. 1987. *The Politics of Program Evaluation.* California, US: SAGE, Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106, —. 1993. "Politics and Evaluation: A Reprise with Mellower Overtones." *American Journal of Evaluation* 14(1):107-09.

[1152] Lipsky, D. B., R. L. Seeber, A. C. Avgar, and R. M. Scanza. 2007. "Managing the Politics of Evaluation: Lessons from the Evaluation of ADR Programs." Pp. 116-29 in *Proceedings of the Fift-Ninth Annual Meeting of the Labor and Employment Relations Association*, Morris, M. 1990. "A Nightmare in Elm City: When Evaluation Field Experiences Meet Organizational Politics." *Evaluation Review* 14(1):91-99.

are related to different stages of this process, i.e. design and commissioning, promotes a more holistic understanding of the relationship between evaluation and politics. In the context of governance, broadening the scope of evaluation as a tool for reducing risk and uncertainty and controlling a more fragmented government structure demands acknowledging the role of all stakeholders involved.

The next section is guided by a more empirical approach. Its purpose is to present a series of practical recommendations for each case study and to outline a possible future research agenda.

7.3 Public Policy Recommendations and Future Research Agendas

An important element of a PhD thesis relies on its capacity to generate new knowledge that can be translated into practical applications in the field of study. This section develops a group of recommendations directed to foster the positive aspects of politics in evaluation and reducing its negative effects on this process. The recommendations presented concern each one of the case studies analysed.

7.3.1 Recommendations for the UK Case

The institutional framework described for the performance of evaluation in the UK (Chapter 4) showed a strong audit system that coexists with a fragmented and irregular evaluation policy. The case studies analysed allowed observing the conditions in which stakeholders interact and the discretionary spaces at their disposal for the pursuit of their interests. In this context, what it is observed is a diminished role of evaluation in policymaking and unclear evidence about its utility. Therefore, in terms of the hidden politics of evaluation, the conditions established by the institutional framework influence the dynamic in which evaluation takes place. In the idea of aligning in a better way evaluation with its political nature, the following recommendations are made:

a) The disconnection between parallel process of assessment (evaluation, audit, and performance measurement) reduces the capability of evaluation to impact

policymaking. Consequently, it is advisable to identify mechanisms of interaction and complementation between these instruments.

b) The hidden politics of evaluation is related to the asymmetries of power and influence among stakeholders. Therefore, more structured mechanisms of participation in these processes could help identifying those stakeholders that are more active and weigh their influence according to their characteristics.

c) Even though there are mechanisms for participating in evaluation/audit processes, these are mostly limited to the provision of evidence. A broader approach to the democratisation of evaluation implies thinking of those key areas of the process that can be enriched by the inclusion of stakeholders, without putting at risk the independence and objectivity of the assessment.

d) Without compromising the necessary freedom required by departments to design and conduct evaluation according to the needs of their programmes, a more structured evaluation policy could reduce the asymmetries existent in terms of the quantity and quality of evaluations performed in central government. In addition, this could generate relevant information about the way in which this activity is undertaken and the areas of improvement.

e) The discussions about the effectiveness of policy should considerate the political values around a particular issue. For example, acknowledging the sensitiveness of a policy issue like cancer can promote a more holistic understanding of how 'success' can be interpreted, considering different views and going beyond the mere interpretation of data.

7.3.2 Recommendations for the Mexican Case

The conditions of the evaluation policy found in Mexico were quite different to the case of the UK. Policy evaluation has been promoted and implemented under a highly centralised scheme controlled by the Federal Government. The evaluation policy has mostly developed a series of instruments and methods to apply in a standardised way to federal programmes, particularly those belonging to the social sector. Furthermore, the audit system operating lacks the political support and technical capacity to effectively hold the government to account. This scenario shows that even though the institutionalisation of evaluation has been strongly promoted, the evaluation policy is still at an early stage in which its effect on policymaking is not very clear. In this context, some recommendations for the Mexican case are exposed:

a) One of the main issues regarding the hidden politics of evaluation is the need for a more developed and open market of evaluators. This could be beneficial for the independence and legitimacy of evaluation and it could avoid negative incentives produced by the closeness between evaluators and commissioners.

b) The lack of flexibility produced by the standardisation of evaluation should be reconsidered in the aim of conducting evaluations that might be more *ad hoc* to the particular needs of programmes.

c) The involvement of external stakeholders could improve the capacity of programmes to conduct more sophisticated and complex studies. For example, the inclusion of NGOs and civil society in the evaluation of the Health Caravans

Programme could generate more specific information about its operation and its impact on the population.

d) In the pursuit of making explicit the independence of evaluation through the use of external consultant, the evaluation policy is excessively centralised and monopolised by the Federal Government. A more open and flexible model could allow a more systematic and incremental incorporation of external views.

e) It is advisable to rethink the role and the faculties given to the SAO; without more institutional mechanisms to hold the government to account, auditing will generate only marginal changes in policy.

These recommendations developed for both cases aim only to highlight some general issues that could be worth a deeper analysis. The implementation of each recommendation is constrained by multiple factors and by the political support of the stakeholders involved. In addition, each one of the recommendations requires a full and detailed development.

From an overall perspective, this thesis has shown that despite the development and adoption of evaluation and audit mechanisms in the public sector that has gained relevance during recent years, the state has not been capable yet of taking full advantage of these tools, particularly in the political sphere. This means that the state can be considered more informed, more open, even more cooperative, but not necessarily smarter as there is still a long way to go in the understanding of the role of evaluation in the context of governance.

7.3.3 Future Research Agendas

The process of undertaking this research has shed new light to new potential areas of research. These issues are mostly related to the political nature of evaluation but also to the relationship of this activity with other mechanisms of assessment. In broad terms, six areas of future research were identified:

1. The relationship between the use of evaluation and the level of public trust in institutions.
2. The complementarity of audit and evaluation approaches for the improvement of public policy.
3. The use evaluation outputs for the construction of political arguments.
4. Evaluation as a mechanism for strengthening the legislative process.
5. The influence of international organisations in policy evaluation and the implications for independence and objectivity.
6. Due to the political reform occurring in Mexico, it would be relevant to analyse the changes in the role of CONEVAL as an autonomous agency in charge of evaluating public policy. Important discussions are taking place in relation to

the consequences of modifying its governance and the implications for its credibility in the political landscape.

In many ways, the discussions presented in this research and the findings obtained through the case studies highlight the relevance of giving more attention to the role of policy evaluation within and outside the governmental sphere. The identification of its utilisation as both a managerial and a political instrument allows reflecting upon multiple future research agendas. Undoubtedly, the hidden politics of evaluation have exposed different potential streams of research about both its implications for academicians and practitioners.

Bibliography

"Act on Acquistion, Leasing and Public Sector Services."

"Agreement that issues the operational rules of the Health Caravans Programme for the 2011 fiscal year."

"Agreement that issues the operational rules of the Integral Quality Health System for the 2014 fiscal year."

"Decree that amends, supplements or repeals certain provisions of the Organic Law of the Federal Public Administration."

"Decree that regulates the National Council for the Evaluation of the Social Development Policy."

"Federal Accountability Law."

"Federal Law of Budget and Fiscal Liability."

"General guidelines for the evaluation of the federal programmes of the Federal Public Administration." Mexico.

"Health and Social Care Act."

"Mechanism for the Follow-up of Aspects Susceptible to Improvement."

"Political Constitution of the United Mexican States."

"Public Administration Modernisation Programme 1995-2000." Mexico City, Mexico.

"The Public Contracts Regulations 2015."

"Social Development General Law." Mexico.

2013a. "Health Caravans registers more than 40, 000 consultations in Veracruz." in *La Crónica*. Mexico City.

2013b. "Osborne's spending review – winners and losers at a glance " *The Guardian*.

2015. "Health Caravans provides 25,600 health consultations." El Correo de Manzanillo

Aberbach, J., and B. Rockman. 2002. "Conducting and Coding Elite Interviews." *Political Science & Politics* 35(4):673-76.

Abma, T. A. 2006. "The Practice and Politics of Responsive Evaluation." *American Journal of Evaluation* 27(1):31-43.

Ackerman, J. M. 2005. "Social Accountability in the Public Sector. A Conceptual Discussion." edited by Social Development Department. Washington, D.C., US: The World Bank.

—. 2007. "Hacia un sistema nacional de tranparencia, fiscalización y rendición de cuentas." in *Finanzas Públicas para el Desarrollo*, edited by J. L. Calva. Mexico City: Miguel Ángel Porrúa.

Acosta, F. 2010. "La evaluación de la política social en México: Avances recientes, tareas pendientes y dilemas persistentes." *Papeles de Población* 64:155-88.

Adams, A. E., N. A. Nnawulezi, and L. Vandenberg. 2014. ""Expectations to Change" (E2C): A Participatory Method for Facilitating Stakeholder Engagement With Evaluation Findings." *American Journal of Evaluation*.

Advisory Committee on the Safety of Blood, Tissues and Organs. 2014. "Transplantation of Organs from Deceased Donors with Cancer or a History of Cancer." Department of Health.

AFP. 2015. "Oliver Sacks describes learning he has terminal cancer: 'My luck has run out' " in *The Telegraph*. London, UK.

Age UK. 2012. "Response to CQC 'State of Care' Report." London, UK.

Albaek, E. 1995. "Policy Evaluation: Design and Utilization." in *Policy Evaluation*, edited by Ray C. Rist. Hants, England: Edward Elgar Publishing Company.

Alcántara, M. A. (Ed.). 2012. *La calidad de la atención a la salud en México a través de sus instituciones: 12 años de experiencia*. Mexico City: Ministry of Health.

Alcántara, S., and H. Jiménez. 2015. "Tibia, actuación de la ASF ante irregularidades: Raya (Lukewarm, SAO performance before inconsistencies: Raya)." in *El Universal*. Mexico City.

Alexander, H. 2003. "Health-Service Evaluations: Should We Expect the Results to Change Practice?" *Evaluation* 9(4):405-14.

Alkin, M. C. 2013. *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences*. California, US: SAGE.

Altschuld, J. W., and H. Y. Zheng. 1995. "Assessing the Effectiveness of Research Organizations: An Examination of Multiple Approaches." *Evaluation Review* 19(2):197-216.

American Evaluation Association. 2007. "AEA Evaluation Policy Task Force Charge."

Andrew G. MP. 2013. "Welcome for Cornwallis Care Quality report."

Andrews, M. 2010. "Good Government Means Different Things in Different Countries." *Governance* 23(1):7-35.

Appleby, J. 2014. "The Cancer Drugs Fund: inequitable and inefficient?". London, UK: The King's Fund

Arellano, D. 2007. "Una cuestión de confianza. Retos e instrumentos para una reforma del Estado en América Latina." *Nueva Sociedad* (210):95-109.

Aristotle. 1995. *Politics*. Oxford, UK: Clarendon.

Arreola, A. (Ed.). 2008. *¿Gobernar por resultados? Implicaciones de la política de evaluación del desempeño del gobierno mexicano.* Mexico City: GESOC, A.C.

ASK Inyathelo. "Advantages and Disadvantages of Internal and External Evaluations."

Aucoin, P. 2012. "New Political Governance in Westminster Systems: Impartial Public Administration and Management Performance at Risk." *Governance* 25(2):177-99.

Azzam, T. 2010. "Evaluator Responsiveness to Stakeholders." *American Journal of Evaluation* 31(1):45-65.

Azzam, T., and B. Levine. 2014. "Negotiating Truth, Beauty, and Justice: A Politically Responsive Approach." *New Directions for Evaluation* 2014(142):57-70.

Babbie, E. 2004. *The Practice of Social Research*. London, UK: Thomson/Wadsworth.

Bache, I., and R. Chapman. 2008. "Democracy through Multilevel Governance? The Implementation of the Structural Funds in South Yorkshire." *Governance: An International Journal of Policy, Administration and Institutions* 21(3):397-418.

Bache, I., L. Reardon, I. Bartle, M. Flinders, and G. Marsden. 2014. "Symbolic Meta-Policy: (Not) Tackling Climate Change in the Transport Sector." *Political Studies* 63(4):830-51.

Bacon, R., and C. Hope. 2013. *Conundrum. Why every government gets things wrong -and what we can do about it*. Great Britain: Biteback Publishing Ltd.

Bailey, J., and P. Paras. 2006. "Perceptions and Attitudes about Corruption and Democracy in Mexico." *Mexican Studies* 22(1):57-82.

Balthasar, A. 2009. "Institutional Design and Utilization of Evaluation: A Contribution to a Theory of Evaluation Influence Based on Swiss Experience." *Evaluation Review* 33(3):226-56.

Bamberger, M. 1991. "The Politics of Policy Evaluation in Developing Countries." *Evaluation and Program Planning* 14(4):325-39.

Bamberger, M., J. Rugh, and L. Mabry. 2012. *Real-World Evaluation: Working Under Budget, Time, Data, and Political Constraints*. California, US: SAGE.

Banner, D. K. 1974. "The politics of evaluation research." *Omega* 2(6):763-74.

Barbier, J. 1999. "Inter-governmental Evaluation: Balancing Stakeholders' Expectations with Enlightenment Objectives?" *Evaluation* 5(4):373-86.

Bardach, E. 1977. *The implementation game: What happens after a bill becomes a law*. Massachusetts, US.

—. 2008. *A Practical Guide For Policy Analysis: The Eightfold Path To More Effective Problem Solving* United States of America: Congressional Quarterly, Inc.

Barrett, S. M. 2004. "Implementation Studies: Time for a Revival? Personal Reflections on 20 Years of Implementation Studies." *Public Administration* 82(2):249-62.

Barrett, S. M., and C. Fudge (Eds.). 1981. *Policy and Action: Essays on the Implementation of Public Policy*. London, UK: Methuen.

Baur, V. E., A. H. G. Van Elteren, C. J. Nierse, and T. A. Abma. 2010. "Dealing with Distrust and Power Dynamics: Asymmetric Relations among Stakeholders in Responsive Evaluation." *Evaluation* 16(3):233-48.

BBC. 2011. "Undercover Care: The Abuse Exposed." London, UK.

BBC News. 2012. "Dame Jo Williams: Care regulator chair to resign." London, UK.

—. 2013. "Morecambe Bay trust failings prompts watchdog criticism." Lancashire, UK.

—. 2014a. "Business Secretary Vince Cable defends Royal Mail sale." London, UK.

—. 2014b. "NHS's cancer help 'postcode lottery' angers Pembrokeshire nurse." London, UK.

—. 2015. "Cancer care in England has lost momentum, say MPs." London, UK.

Bechar, S., and I. Mero-Jaffe. 2014. "Who Is Afraid of Evaluation? Ethics in Evaluation Research as a Way to Cope With Excessive Evaluation Anxiety: Insights From a Case Study." *American Journal of Evaluation* 35(3):364-76.

Behn, R. D. 2003. "Why Measure Performance? Different Purposes Require Different Measures." *Public Administration Review* 63(5):586-606.

Behrman, J. R., and E. Skoufias. 2006. "Mitigating Myths about Policy Effectiveness: Evaluation of Mexico's Antipoverty and Human Resource Investment Program." *The ANNALS of the American Academy of Political and Social Science* 606(1):244-75.

Bemelmans-Videc, M. L., R. C. Rist, and E. Vedung. 2003. *Carrots, Sticks & Sermons: Policy Instruments and their Evaluation*. London UK: Transaction Publishers.

Benjamin, L. M. 2008. "Evaluator's Role in Accountability Relationships: Measurement Technician, Capacity Builder or Risk Manager?" *Evaluation* 14(3):323-43.

Benton, M., and M. Russell. 2013. "Assessing the Impact of Parliamentary Oversight Committees: The Select Committees in the British House of Commons." *Parliamentary Affairs* 66(4):772-97.

Benwell, R., and O. Gay. 2011. "The Separation of Powers." London, UK.

Berriet-Solliec, M., P. Labarthe, and C. Laurent. 2014. "Goals of evaluation and types of evidence." *Evaluation* 20(2):195-213.

Bevir, M., R. A. W. Rhodes, and P. Weller. 2003. "Traditions of Governance: Interpreting the Changing Role of the Public Sector." *Public Administration* 81(1):1-17.

Birbeck University of London. 2012. "National Evaluation of Sure Start." London, UK.

Birkland, T. A. 2011. *An Introduction to the Policy Process: Theories, Concepts, and Models of Public Policy Making*. New York, US: M.E. Sharpe.

Bjørnholt, B., and F. Larsen. 2014. "The politics of performance measurement: 'Evaluation use as mediator for politics'." *Evaluation* 20(4):400-11.

Blomberg, T. G., and G. P. Waldo. 2002. "Evaluation Research, Policy, and Politics." *Evaluation Review* 26(3):340-51.

Boardman, C. 2014. "Assessing Governance: The Importance of Evaluating Policy Outcomes in National Mission Areas." *Governance* 27(3):519-26.

Bogenschneider, K., and T. J. Corbett. 2010. *Evidence-Based Policymaking: Insights from Policy-Minded Researchers and Research-Minded Policymakers*. New York, US: Routledge.

Borland, S. 2012. "Care quango chief is forced to quit over a catalogue of failures (but she walks away with a £1.35m pension pot)." in *The Daily Mail*. London, UK.

Boseley, S. "Half of people in Britain born after 1960 will get cancer, study shows " in *The Guardian*. London, UK.

Bostock, N. 2015. "Cancer care slipping under coalition government, MPs warn." London, UK.

Bourn, J. 2007. *Public Sector Auditing: Is it value for money?* Chichester, UK: John Wiley & Sons, Ltd.

Bovaird, T. 2005. "Public governance: balancing stakeholder power in a network society." *International Review of Administrative Sciences* 71(2):217-28.

Bovens, M. 2007. "Analysing and Assessing Accountability: A Conceptual Framework." *European Law Journal* 13(4):447-68.

Bovens, M., P. 't Hart, and S. Kuipers. 2006. "The Politics of Policy Evaluation." in *The Oxford Handbook of Public Policy*, edited by M. Moran, M. Rein, and R. E. Goodin. Oxford, UK: Oxford University Press.

Bovens, M., R.E. Goodin, and T. Schillemans (Eds.). 2014. *The Oxford Handbook of Public Accountability*. Oxford, UK: Oxford University Press.

Boyer, J. F., and L. I. Langbein. 1991. "Factors Influencing the Use of Health Evaluation Research in Congress." *Evaluation Review* 15(5):507-32.

Bozeman, B. 2004. "The Internet's Impact on Policy Evaluation: Information Compression and Credibility." *Evaluation Review* 28(2):156-74.

Brady, E. 2014. "Labour Pledges One-Week Cancer Diagnosis." London, UK: Sky News.

Brandon, P. R., and J. M. Singh. 2009. "The Strength of the Methodological Warrants for the Findings of Research on Program Evaluation Use." *American Journal of Evaluation* 30(2):123-57.

Breast Cancer Care. "Facts and statistics."

—. 2015. "The Breast Cancer Manifesto 2015. Breast Cancer Care's ambitions for the next government." London, UK.

Brewer, R. 2007. *Your PhD Thesis. How to plan, draft, revise and edit your thesis*. Abergele, UK: Studymates Limited.

British Prime Minister's Office. 2010. "Business Plans: track progress in implementing our policies." London, UK.

Brooks, R. A. 1996. "Blending two cultures: State legislative auditing and evaluation." *New Directions for Evaluation* 1996(71):15-28.

Brown, A. B., and J. A. Klerman. 2012. "Independent Evaluation: Insights from Public Accounting." *Evaluation Review* 36(3):186-219.

Bryman, A. 1984. "The Debate about Quantitative and Qualitative Research: A Question of Method or Epistemology?" *The British Journal of Sociology* 35(1):75-92.

—. 2012. *Social Research Methods*. Oxford, UK: Oxford University Press.

Burnham, P., K. Gilland Lutz, W. Grant, and Z. Layton-Henry. 2008. *Research Methods in Politics*. Basingstoke, UK: Palgrave Macmillan.

Burton, M. 2013. "The Politics of Public Sector Reform." Hampshire, UK: Palgrave Macmillan.

Cabinet Office. 1999a. "Modernising Government." London, UK.

—. 1999c. "Professional Policy Making for the Twenty First Century."

—. 2000. "Wiring it up: Whitehall's Management of Cross Cutting Politics and Services." London, UK.

—. 2013a. "Public bodies reform." London, UK.

—. 2013b. "What Works: evidence centres for social policy." London, UK.

—. 2014a. "CBI Annual Conference 2014: Prime Minister's address."

—. 2014b. "Giving Evidence to Select Committees. Guidance for Civil Servants."

Cabrero, E., and G. Nava (Eds.). 1999. *Gerencia Pública Municipal: Conceptos Básicos y Estudios de Caso*. Mexico City: Miguel Ángel Porrúa-CIDE.

Campbell, D. 2011. "Care Quality Commission's leaders 'lack the necessary skills' " in *The Guardian*. London, UK.

—. 2012. "Care Quality Commission problems-timeline." in *The Guardian*. London, UK.

—. 2014a. "Almost half of cancer patients diagnosed too late." in *The Guardian*. London, UK.

—. 2014b. "Cancer patients dying early because of postcode lottery " in *The Guardian*. London, UK.

—. 2015. "Government accused of breaking promises on NHS cancer treatment " in *The Guardian*. London, UK.

Campbell, D., S. Redman, L. Jorm, M. Cooke, A. Zwi, and L. Rychetnik. 2009. "Increasing the use of evidence in health policy: Practice and views of

policy makers and researchers." *Australia and New Zealand Health Policy* 6(1):21.

Campbell, D. T., and J. C. Stanley. 1966. *Experimental and Quasi-Experimental Designs for Research*. Chicago, US: Houghton Mifflin.

Cancer Research UK. "Ninety per cent approve of cancer screening but screening uptake is lower." London, UK.

—. 2013. "Cancer Statistics Key Facts." London, UK.

—. 2014a. "Cancer Statistics Report: Cancer Mortality in the UK in 2012."

—. 2014b. "Every Moment Counts. Our Annual Review 2013/14.". London, UK: Cancer Research UK.

—. 2015a. "£14 million Government investment to help 'personalise' treatment." London, UK.

—. 2015b. "Cancer Drugs Fund." London, UK.

—. 2015d. "Standard cigarette packs 'could save 2000 lives a year'." London, UK.

—. 2015e. "What is cancer?". London, UK.

—. 2015f. "Written evidence from Cancer Research UK." London, UK.

Cardozo, M. 2009. "La institucionalización de una cultura de la evaluación en la administración pública mexicana: Avances y desafíos pendientes." *Convergencia. Revista de Ciencias Sociales* 49:175-98.

Care Home. 2015. "CQC announce nine per cent increase in regulatory fees." London, UK.

Care Quality Commission. 2011a. "Care Quality Commission response to the National Audit Office Report."

—. 2011b. "Dignity and Nutrition Inspection Programme. National Overview.". London, UK.

—. 2012a. "Performance and Capability Review. CQC Action Plan." London, UK.

—. 2012b. "The state of health care and adult social care in England ". London, UK.

—. 2013a. "Annual report and accounts 2012/13." London, UK.

—. 2013c. "CQC Corporate Governance Framework." London, UK.

—. 2013d. "Framework Agreement between the Department of Health and Care Quality Commission." London, UK.

—. 2013g. "A new start. Consultation on the changes to the way the CQC regulates, inspects and monitors care." London, UK.

—. 2013i. "Raising standards, putting people first. Our strategy for 2013-2016.". London, UK.

—. 2013k. "Working for CQC." London, UK.

—. 2014a. "Business Plan 2014/15 to 2015/16." London, UK.

—. 2014b. "CQC welcomes new Board members." London, UK.

—. 2014c. "Whipps Cross University Hospital. Quality Report." London, UK.

—. 2015a. "Board Members." London, UK.

—. 2015b. "Spending transparency." London, UK.

—. 2015c. "What we do on an inspection." London, UK.

Carman, J. G. 2009. "Nonprofits, Funders, and Evaluation: Accountability in Action." *The American Review of Public Administration* 39(4):374-90.

Castillo, M. A. 2015. "Gestión y evaluación: De la resistencia al desarrollo próximo." in *Segundo Congreso Latinoamericano de Medición y Evaluación Educacional*. Mexico City.

Castro, M. F., G. Lopez-Acevedo, G. Beker Busjeet, and X. Fernandez-Ordonez. 2009. "Mexico's M&E System: Scaling Up from the Sectoral to the National Level." Washington, D.C., US: The World Bank.

Cejudo, G. 2011. "De las recomendaciones a las acciones: El uso de la evaluación en la mejora del desempeño del gobierno." in *De las recomendaciones a las acciones: La experiencia del Premio 2011 Programas Federales Comprometidos con el Proceso de Evaluación*, edited by G. Cejudo and C. Maldonado. Mexico City: CIDE-SFP.

Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS). 2010. "Indigenous Communication Programme Evaluation." Mexico City.

Chancellor of the Exchequer. 1998. "Modern Public Services for Britain: Investing in Reform. Comprehensive Spending Review: New Public Spending Plans 1999-2002."

Chantler, C. 2014. "Standardised packaging of tobacco." London, UK: King's College London.

Chaplowe, S. G., and R. B. Engo-Tjéga. 2007. "Civil Society Organizations and Evaluation: Lessons from Africa." *Evaluation* 13(2):257-74.

Chávez, V. 2014. "Pleitos ponen en jaque al Ifai; ASF sin dientes (Lawsuits threatens IFAI; SAO toothless)." in *El Financiero* Mexico City.

—. 2015. "En 14 años, la ASF sólo ha castigado a siete personas (In 14 years, the SAO has only punished seven persons)." in *El Financiero*. Mexico City.

Chelimsky, E. 1985. "Comparing and Contrasting Auditing and Evaluation: Some Notes on their Relationship." *Evaluation Review* 9(4):483-503.

—. 1987. "What Have We Learned about the Politics of Program Evaluation?" *Educational Evaluation and Policy Analysis* 9(3):199-213.

—. 1995. "The Political Environment of Evaluation and What it Means for the Development of the Field: Evaluation for a New Century: A Global Perspective." *American Journal of Evaluation* 16(3):215-25.

—. 1996. "Auditing and Evaluation: Whither the Relationship?" *New Directions for Evaluation* 1996(71):61-67.

—. 2007. "Factors Influencing the Choice of Methods in Federal Evaluation Practice." *New Directions for Evaluation* 2007(113):13-33.

—. 2008. "A Clash of Cultures: Improving the "Fit" Between Evaluative Independence and the Political Requirements of a Democratic Society." *American Journal of Evaluation* 29(4):400-15.

—. 2009. "Integrating Evaluation Units into the Political Environment of Government: The Role of Evaluation Policy." *New Directions for Evaluation* 123:51-66.

—. 2012. "Valuing, Evaluation Methods, and the Politicization of the Evaluation Process." *New Directions for Evaluation* 2012(133):77-83.

Chen, H. T. 2005. *Practical Program Evaluation: Assessing and Improving Planning, Implementation, and Effectiveness*. London, UK: SAGE.

Chouinard, J. A. 2013. "The Case for Participatory Evaluation in an Era of Accountability." *American Journal of Evaluation* 34(2):237-53.

Clarke, J. 2005. "Performing for the Public: Doubt, Desire, and the Evaluation of Public Services." in *The Values of Bureaucracy*, edited by P. Du Gay. Oxford, UK: Oxford University Press.

—. 2008. "Performance Paradoxes: The Politics of Evaluation in Public Services." Pp. 120-34 in *Public Services Inspection in the UK. Research Highlight in Social Work*, edited by H. Davis and S. Martin. London, UK: Jessica Kingsley Publishers.

Coffey International Development. 2014. "Scoping the Need for Improved Access to Funding, Advisory Support and Business Skills Development for SMEs in Afghanistan." London, UK.

Committee of Public Accounts. 2011. "Delivering the Cancer Reform Strategy." London, UK.

—. 2012. "The Care Quality Commission: Regulating the quality and safety of health and adult social care." London, UK.

—. 2013a. "Select Committees. Government Responses." London, UK.

—. 2013b. "Tax avoidance: The role of large accountancy firms ". London, UK: House of Commons.

—. 2014a. "Adult social care in England." London, UK: House of Commons.

—. 2014c. "Publications." London, UK.

—. 2015a. "Chair's statement on improving cancer services and outcomes in England."

—. 2015b. "List of Reports from the Committee during the current Parliament." London, UK: House of Commons.

—. 2015c. "Oral evidence: Progress in improving cancer services and outcomes." London, UK: House of Commons.

Conley-Tyler, M. 2005. "A fundamental choice: Internal or external evaluation?" *Evaluation Journal of Australasia* 4(1 & 2):3-11.

Considine, M. 2002. "The End of the Line? Accountable Governance in the Age of Networks, Partnerships, and Joined-Up Services." *Governance: An International Journal of Policy, Administration, and Institutions* 15(1):21-40.

Contandriopoulos, A. P. 1999. "Is the institutionalization of evaluation sufficient to guarantee its practice?" *Cadernos de Saúde Pública* 15:253-56.

Contandriopoulos, D., and A. Brousselle. 2012. "Evaluation models and evaluation use." *Evaluation* 18(1):61-77.

Contracts Finder. 2015. "The Provission of A & E Qualitative Research." London, UK.

Cook, J. R. 2014. "Using Evaluation to Effect Social Change: Looking Through a Community Psychology Lens." *American Journal of Evaluation* 36(1):107-17.

Cooper, C. 2015. "Cancer Drugs Fund: Life-extending drugs to be denied to NHS patients in England as fund overspends " in *The Independent*. London, UK.

Coordinación Nacional del Programa de Desarrollo Humano Oportunidades. 2007. "Agenda Integral de Evaluación del Programa Oportunidades 2007 - 2012 (Evaluation agenda for the *Oportunidades* Programme 2007-2012)." Mexico City: SEDESOL.

Copestake, J. 2014. "Credible impact evaluation in complex contexts: Confirmatory and exploratory approaches." *Evaluation* 20(4):412-27.

Creswell, J. W. 1994. *Research Design: Qualitative and Quantitative Approaches*. California, US.: SAGE.

Cunill-Grau, N., and S. M. Ospina. 2012. "Performance measurement and evaluation systems: Institutionalizing accountability for governmental

results in Latin America." *New Directions for Evaluation* 2012(134):77-91.

Curristine, T. 2005. "Performance Information in the Budget Process: Results of the OECD 2005 Questionnaire." *OECD Journal on Budgeting* 5(2):87-131.

Curristine, T., Z. Lonti, and I. Joumard. 2007. "Improving Public Sector Efficiency: Challenges and Opportunities " *OECD Journal on Budgeting* 7(1):1-42.

Curtis, P., and H. Mulholland. 2011. "Panorama care home abuse investigation prompts government review " in *The Guardian*. London, UK.

D'Arcy, M. 2015. "Cracks in the PAC?". London, UK: BBC News.

Dahl, R. A. 1957. "The concept of power." *Behavioral Science* 2(3):201-15.

Daniell, K. A., A. Morton, and D. Ríos Insua. 2015. "Policy Analysis and Policy Analytics." *Annals of Operations Research*.

Datta, L. E. 1999. "The ethics of evaluation neutrality and advocacy." *New Directions for Evaluation* 1999(82):77-88.

—. 2000. "Seriously Seeking Fairness: Strategies for Crafting Non-partisan Evaluations in a Partisan World " *American Journal of Evaluation* 21(1):1-14.

—. 2001. "Coming attractions." *The American Journal of Evaluation* 22(3):403-08.

—. 2009. "Golden Is the Sand: Memory and Hope in Evaluation Policy and Wvaluation Practice." *New Directions for Evaluation* 123:33-50.

—. 2011. "Politics and Evaluation: More Than Methodology." *American Journal of Evaluation* 32(2):273-94.

Davies, A., C. Ariti, T. Georghiou, and M. Bardsley. 2015. "Evaluation of complex health and care interventions using retrospective matched control methods." Nuffield Trust.

Davies, P. 2012. "The State of Evidence-Based Policy Evaluation and its Role in Policy Formation." *National Institute Economic Review* 219(1):R41-R52.

Davies, R., D. Randall, and R. E. West. 2015. "Using Open Badges to Certify Practicing Evaluators." *American Journal of Evaluation* 36(2):151-63.

De Lay, P., and V. Manda. 2004. "Politics of Monitoring and Evaluation: Lessons from the AIDS Epidemic." *New Directions for Evaluation* 2004(103):13-31.

DeLeon, P. 1987. "Policy Termination as a Political Phenomenon." in *The Politics of Program Evaluation*, edited by Dennis J. Palumbo. California, US: SAGE.

Denzin, N. K. 2009. "The elephant in the living room: or extending the conversation about the politics of evidence." *Qualitative Research* 9(2):139-60.

Department for Business, Innovation and Skills. 2014. "Impact assessment template for government policies." London, UK.

Department for Communities and Local Government. 2015a. "2010 to 2015 Government Policy: House Building." London, UK.

—. 2015b. "Local audit framework replacing the Audit Commission." London, UK.

Department for Culture, Media and Sport. 2013. "An Independent Review of E-Lending in Public Libraries in England." London, UK.

Department for Education. 2011. "National evaluation of Sure Start local programmes: an economic perspective." London, UK.

Department for International Development. 2010. "Joint Irish Aid and DFID's country programme evaluation Tanzania 2004/05 to 2009/10." London, UK.

—. 2013. "International Development Evaluation Policy." London, UK.

—. 2014. "Evaluations completed: April 2014 to March 2015." London, UK.

Department of Health. 2006a. "Autumn Performance Report." London, UK.

—. 2006b. "Best Research for Best Health." London, UK.

—. 2007. "Cancer Reform Strategy." edited by Department of Health. London, UK.

—. 2010. "Business Plan." London, UK.

—. 2011a. "Evaluation of nurse and pharmacist independent prescribing." London, UK.

—. 2011b. "Improving Outcomes: A Strategy for Cancer." London, UK.

—. 2011e. "Improving Outcomes: A Strategy for Cancer. First Annual Report." London, UK.

—. 2011f. "The likely impact of earlier diagnosis of cancer on costs and benefits to the NHS." London,UK.

—. 2012a. "DH business plan quarterly data summary: January 2012." London, UK.

—. 2012b. "Improving Outcomes: A Strategy for Cancer. Second Annual Report 2012." London, UK.

—. 2012c. "Performance and Capability Review. Care Quality Commission." London, UK.

—. 2013a. "Department of Health 2013 Stakeholder Research." London, UK.

—. 2013b. "Health and social care outcomes frameworks." London, UK.

—. 2013c. "Helping more people survive cancer." London, UK.

—. 2013d. "Improving Outcomes: A Strategy for Cancer. Third Annual Report." London, UK.

—. 2014a. "DH Corporate Plan 2014- 2015." edited by Department of Health. London, UK.

—. 2014b. "Mid-year report 2014-15." London, UK.

—. 2014c. "Publications: Independent Reports." London, UK.

—. 2015a. "Government backs standarised packaging of tobacco."

—. 2015b. "Safer care for patients." London, UK.

Department of Health, and NHS England. 2012. "Quality of Life of Cancer Survivors in England. Report on a pilot survey using Patient Reported Outcome Measures (PROMS)." edited by Department of Health. London, UK.

Department of Health, Public Health England, and NHS England. 2014. "Improving Outcomes: A Strategy for Cancer. Fourth Annual Report." London, UK.

Desmond, M. 2004. "Methodological challenges posed in studying an elite in the field." *Area* 36(3):262-69.

Dilanthi, A., and D. Baldry. 2001. "Case study methodology as a means of theory building: performance measurement in facilities management organisations." *Work Study* 50(3):95-105.

Dixit, A. 2002. "Incentives and Organizations in the Public Sector: An Interpretative Review " *Journal of Human Resources* 37(4):696-727.

Dolowitz, D. P., and D. Marsh. 1996. "Who Learns What from Whom: a Review of the Policy Transfer Literature." *Political Studies* 44(2):343-57.

—. 2000. "Learning from Abroad: The Role of Policy Transfer in Contemporary Policy-Making." *Governance* 13(1):5-23.

Donaldson, S. I. 2005. "Using program theory–driven evaluation science to crack the Da Vinci code." *New Directions for Evaluation* 2005(106):65-84.

Donaldson, S. I., and M. Scriven. 2003. *Evaluating Social Programs and Problems: Visions for the New Millennium*. London, UK: Erlbaum.

Dorey, P. 2014. *Policy Making in Britain: An Introduction*. London, UK: SAGE.

Dorotinsky, W. 2011. *Implementing Results Based Budgeting in Thailand*. Washington D.C., US: The World Bank.

Dreaper, J. 2013. "CQC whistle-blower: 'I am tainted because I spoke up'." BBC News.

du Toit, A. 2012. "Making Sense of 'Evidence'. Notes on the Discursive Politics of Research and Pro-Poor Policy Making." in *Working Paper 21*. Bellville: PLAAS, UWC.

Dubnick, M. J. 2002. "Seeking Salvation for Accountability." in *Annual Meeting of the American Political Science Association*.

—. 2005. "Accountability and the Promise of Performance: In Search of the Mechanisms." *Public Performance & Management Review* 28(3):376-417.

—. 2011. "More over Daniel: We need some 'accountability space'." *Administration & Society* 43(6):704-16.

Dubnick, M. J., and H. G. Frederickson. 2011a. *Accountable Governance: Problems and Promises*. London: Routledge.

—. 2011b. "Public Accountability: Performance Measurement, the Extended State, and the Search for Trust." Washington, D.C., US: Kettering Foundation.

Duit, A., and V. Galaz. 2008. "Governance and Complexity—Emerging Issues for Governance Theory." *Governance* 21(3):311-35.

Dunleavy, P., C. Gilson, S. Bastow, and J. Tinkler. 2009. "The National Audit Office, the Public Accounts Committee and the risk landscape in UK public policy." London, UK: The Risk and Regulation Advisory Council.

Dunworth, T., J. Hannaway, J. Holahan, and M. Austin Turner. 2008. "Beyond Ideology, Politics and Guesswork: The Case for Evidence-Based Policy." edited by Urban Institute.

Eddy, R. M., and T. Berry. 2009. "The Evaluator's Role in Recommending Program Closure." *American Journal of Evaluation* 30(3):363-76.

Ellis, K. 2011. "'Street-level Bureaucracy' Revisited: The Changing Face of Frontline Discretion in Adult Social Care in England." *Social Policy & Administration* 45(3):221-44.

Elston, T. 2012. "Developments in UK executive agencies: Re-examining the 'disaggregation–reaggregation' thesis." *Public Policy and Administration* 28(1):66-89.

Estrada García, K. 2015. "Health Caravans come to Ascensión." in *Diario.mx*. Chihuahua, Mexico.

Estrich, S. 2013. "The Politics of Cancer."

European Commission. 2015. "Evidence-based Policymaking."

Evans, M., and J. Davies. 1999. "Understanding Policy Transfer: A Multi-Level, Multi-Disciplinary Perspective." *Public Administration* 77(2):361-85.

Faguet, G. B. 2008. *The war on cancer: An anatomy of failure, a blueprint for the future*. London, UK: Springer.

Farr, J., J. S. Hacker, and N. Kazee. 2008. "Revisiting Lasswell." *Policy Sciences* 41(1):21-32.

Faulkner, W. N. 2014. "A critical analysis of a randomized controlled trial evaluation in Mexico: Norm, mistake or exemplar?" *Evaluation* 20(2):230-43.

Fawcett, P., and D. Marsh. 2014. "Depoliticisation, governance and political participation." *Policy & Politics* 42(2):171-88.

Feeney, M. 2012. "Leaving Office Calderón Leaves Bloody Drug War Legacy, Tens of Thousands Dead and Missing." in *Hit and Run*.

Feinstein, O., and G. Hernández. 2008. "El Papel de la evaluación en México." Banco Mundial.

Feinstein, O. N. 2002. "Use of Evaluations and the Evaluation of their Use." *Evaluation* 8(4):433-39.

Finlayson, A. 1998. "Tony Blair and the jargon of modernisation." *Soundings Autumn* (10).

—. 2012. "Cameron, Culture and the Creative Class: The Big Society and the Post-Bureaucratic Age." *The Political Quarterly* 82(S1):35-47.

Fitzpatrick, J. L. 1989. "The Politics of Evaluation With Privatized Programs: Who Is the Audience?" *Evaluation Review* 13(6):563-78.

Fleischer, D. N., and C. A. Christie. 2009. "Evaluation Use. Results From a Survey of U.S. American Evaluation Association Members." *American Journal of Evaluation* 30(2):158-75.

Fletcher, G., and S. Dyson. 2013. "Evaluation as work in progress: Stories of shared learning and development." *Evaluation* 19(4):419-30.

Flick, U. 2014. *An Introduction to Qualitative Research*. California, US: SAGE.

Flinders, M. 2002. "Governance in Whitehall." *Public Administration* 80(1):51-75.

—. 2008. "The Future of the State." *The Political Quarterly* 79(S1):19-40.

— (Ed.). 2009. *The Oxford Handbook of British Politics*. Oxford, UK: Oxford University Press.

—. 2011. "Daring to be a Daniel: The Pathology of Politicized Accountability in a Monitory Democracy." *Administration & Society* 43(5):595-619.

Flinders, M., and J. Buller. 2006. "Depoliticisation: Principles, Tactics and Tools." *British Politics* 1(3):293-318.

Flyvbjerg, B. 2006. "Five Misunderstandings About Case-Study Research." *Qualitative Inquiry* 12(2):219-45.

Forss, K., and J. Carlsson. 1997. "The Quest for Quality—Or Can Evaluation Findings Be Trusted?" *Evaluation* 3(4):481-501.

Foss Hansen, H. 2005. "Choosing Evaluation Models: A Discussion on Evaluation Design." *Evaluation* 11(4):447-62.

Frederiksen, L. F., F. Hansson, and S. B. Wenneberg. 2003. "The Agora and the Role of Research Evaluation." *Evaluation* 9(2):149-72.

Freeguard, G., P. Bouchal, R. Munro, C. Nimmo, and J. McCrae. 2014. "Whitehall Monitor 2014. A data-driven analysis of the size, shape and performance of Whitehall." London, UK: Institute for Government.

Fukuyama, F. 2013. "What Is Governance?" *Governance* 26(3):347-68.

Furubo, J., and O. Karlsson. 2011. "Evaluation: For Public Good or Professional Power?" Pp. 240 in *Evaluation: Seeking Truth or Power?*, edited by P. Eliadis, J. E. Furubo, and S. Jacob. New Jersey, US: Transaction Publishers.

Gaarder, M. M., and B. Briceño. 2010. "Institutionalisation of Government Evaluation: Balancing Trade-Offs." International Initiative for Impact Evaluation (3ie).

García López, R. , and M. García Moreno. 2011. "Managing for Development Results. Progress and Challenges in Latin America and the Caribbean." Washington, D.C., US: Inter-American Development Bank.

Gay, O. 2005. "Public Service Agreements." Parliament and Constitution Centre.

Gerring, J. 2004. "What Is a Case Study and What Is It Good for?" *American Political Science Review* 98(2):341-54.

Gerston, L. N. 2010. *Public Policy Making: Process and Principles*. New York, US: M.E. Sharpe.

Gertler, P. J., S. Martínez, P. Premand, L. B. Rawlings, and C. M. J. Vermeesch. 2011. "La evaluación de impacto en la práctica." Washington, D.C., US: Banco Internacional de Reconstrucción y Fomento/Banco Mundial.

GESOC, A.C. 2009. "Monitoreo de la Confiabilidad de las Evaluaciones a Programas Federales (2007-2008)."

—. 2014. "Índice de Desempeño de Programas Públicos Federales 2014 (Performance Index of Federal Programmes 2014)." Mexico City.

Geva-May, I. 2004. "Riding the Wave of Opportunity: Termination in Public Policy." *Journal of Public Administration Research and Theory: J-PART* 14(3):309-33.

Geva-May, I., and W. Thorngate. 2003. "Reducing Anxiety and Resistance in Policy and Programme Evaluations. A Socio-Psychological Analysis." *Evaluation* 9(2):205-27.

Gildemyn, M. 2014. "Understanding the Influence of Independent Civil Society Monitoring and Evaluation at the District Level: A Case Study of Ghana." *American Journal of Evaluation* 35(4):507-24.

Ginsburg, A., and N. Rhett. 2003. "Building a Better Body of Evidence: New Opportunities to Strengthen Evaluation Utilization." *American Journal of Evaluation* 24(4):489-98.

Global Democracy Ranking. 2014. "The Democracy Ranking of the Quality of Democracy."

Gobierno Federal. 2012. "Sexto Informe de Gobierno."

Goertz, G. 2006. "Introduction to the Special Issue "Causal Complexity and Qualitative Methods"." *Political Analysis* 14(3):223-26.

Goldacre, B. 2013. "Building evidence into education." London, UK.

Gomm, R., M. Hammersley, and P. Foster (Eds.). 2000. *Case Study Method: Key Issues, Key Texts*. London, UK: SAGE.

González Arredondo, A. (Ed.). 2008. *¿Gobernar por resultados? Implicaciones de la política de evaluación del desempeño del gobierno mexicano*. Mexico City: GESOC, A.C.

Government of the State of Chiapas. 2013. "We take health where it is needed: Manuel Velasco." Chiapas, Mexico.

Government Social Research Service. "Professional Guidance." London, UK.

Gray, A. , and B. Jenkins. 1993. "Separate Developments: Budgeting, Auditing, and Evaluation in the United Kingdom." in *Budgeting, Auditing, and*

Evaluation. Functions and Integration in Seven Governments edited by A. Gray, B. Jenkins, and B. Segsworth. London, UK: Transaction Publishers.

Gray, A., and B. Jenkins. 2011. "Policy and Evaluation: Many Powers, Many Truths." in *Evaluation. Seeking Truth or Power?*, edited by P. Eliadis, J. E. Furubo, and S. Jacob. New Jersey, US: Transaction Publishers.

Green-Pedersen, C., and S. Walgrave (Eds.). 2014. *Agenda Setting, Policies, and Political Systems: A Comparative Approach*. Chicago, US: The University of Chicago Press.

Green, L. W., J. M. Ottoson, C. Garcia, and R. A. Hiatt. 2009. "Diffusion theory and knowledge dissemination, utilization, and integration in public health." *Annu Rev Public Health* 30:151-74.

Greene, J. C. 1997. "Evaluation as Advocacy." *American Journal of Evaluation* 18(1):25-35.

—. 2001. "Dialogue in Evaluation: A Relational Perspective." *Evaluation* 7(2):181-87.

—. 2003. "War and peace … and evaluation." *Studies in Educational Policy and Educational Philosophy* 2.

—. 2007. *Mixed Methods in Social Inquiry*. California, US: Jossey-Bass.

Greene, J. C., L. DeStefano, H. Burgon, and J. Hall. 2006. "An educative, values-engaged approach to evaluating STEM educational programs." *New Directions for Evaluation* 2006(109):53-71.

Greener, I., and B. Greve (Eds.). 2014. *Evidence and Evaluation in Social Policy*. London, UK: Wiley Blackwell.

Greenhalgh, T., and J. Russell. 2009. "Evidence-Based Policymaking: A Critique." *Perspectives in Biology and Medicine* 52(2):304-18.

Greig, R. 2013. "CQC – 3 Big Reasons It Went Wrong; 3 Big Actions to Get It Right." London, UK: NDTi National Development Team for inclusion.

Greve, C., M. Flinders, and S. Van Thiel. 1999. "Quangos—What's in a Name? Defining Quangos from a Comparative Perspective." *Governance* 12(2):129-46.

Griffiths, S., S. Ahmar, J. Seager, L. Simbayi, and N. Mbelle. 2014. "Mid-Term Evaluation of the SARRAH programme." Department for International Development.

Gris Legorreta, P. C. 2011. "La evaluación de políticas públicas como instrumento para la rendición de cuentas." *Trimestre Fiscal* 99:89-103.

—. 2013. "Auditoría y Evaluación: ¿Hacia un tercer enfoque?" *XIII Certamen Nacional de Ensayo Sobre Fiscalización y Rendición de Cuentas*.

Grob, G. F. 2014. "How to Become an Effective Advocate Without Selling Your Soul." *American Journal of Evaluation* 35(3):391-97.

Gudmundsson, H. 2003. "The Policy Use of Environmental Indicators - Learning from Evaluation Research." *The Journal of Transdisciplinary Environmental Studies* 2(2):1-12.

Guenther, J. 2008. "The Role of the Evaluator in Remote Contexts." in *Public Seminar*. Alice Springs: Centre for Remote Health.

Guenther, J., E. Williams, and A. Arnott. 2010. "The politics of evaluation: evidence-based policy or policy-based evidence?" in *NARU Public Seminar Series*. Darwin, Australia.

Guerrero, J. P. 1995. "La evaluación de las políticas públicas: enfoques teóricos y realidades en nueve países desarrollados." *Gestión y Política Pública* IV(1):47-115.

Hague, R., M. Harrop, and S. Breslin. 1992. *Comparative Government and Politics*. London, UK: The McMillan Press Ltd.

Hanberger, A. 2006. "Evaluation of and for Democracy." *Evaluation* 12(1):17-37.

—. 2011. "The real function of evaluation and response systems."

Hans Klijn, E., and J. F.M. Koppenjan. 2014. "Accountable Networks." in *The Oxford Handbook of Public Accountability*, edited by M. Bovens, R. E. Goodin, and T. Schillemans. Oxford, UK: Oxford University Press.

Hansson, F. 2006. "Organizational Use of Evaluations: Governance and Control in Research Evaluation." *Evaluation* 12(2):159-78.

Hansson, F., M. T Norn, and T. Bungaard Vad. 2014. "Modernize the public sector through innovation? A challenge for the role of applied social science and evaluation." *Evaluation* 20(2):244-60.

Harrison, S., D. J. Hunter, and C. Pollitt. 1990. *The dynamics of British health policy*. London,UK: Unwin Hyman.

Hasenfeld, Y., K. Hill, and D. Weaver. "A Participatory Model for Evaluating Social Programs." San Francisco, CA: The James Irvine Foundation.

Hatry, H. P., J. S. Wholey, and K. E. Newcomer. 2010. "Evaluation Challenges, Issues and Trends." in *Handbook of Practical Program Evaluation*, edited by Harry P. Hatry, Joseph S. Wholey, and Kathryn E. Newcomer. US: Jossey-Bass.

Haveri, A. 2008. "Evaluation of Change in Local Governance: The Rhetorical Wall and the Politics of Images." *Evaluation* 14(2):141-55.

Hay, C. 2007. *Why we hate politics*. Cambridge, UK: Polity Press.

—. 2014. "Depoliticisation as process, governance as practice: What did the 'first wave' get wrong and do we need a 'second wave' to put it right?" *Policy & Politics* 42(2):293-311.

Heady, F. 2001. *Public Administration: A Comparative Perspective*. New York, US: Marcel Dekker.

Health Committee. 2011a. "Accountability hearing with the Care Quality Commission." London, UK: House of Commons.

—. 2011c. "Annual accountability hearing with the Care Quality Commission." London, UK.

—. 2012. "2012 Accountability Hearing with the Care Quality Commission." London, UK.

—. 2013a. "2013 Annual accountability hearing with the Care Quality Commission." London, UK: House of Commons.

—. 2013b. "Written evidence from the Relatives & Residents Association ". London, UK.

—. 2014a. "2013 Accountability Hearing with the Care Quality Commission." London, UK.

—. 2014d. "2014 Accountability hearing with the Care Quality Commission." London, UK: House of Commons.

Health Management and Policy Alert. 2015. "Progress In Improving Cancer Services And Outcomes In England." London, UK.

Hedrick, T. E. 1988. "The Interaction of Politics and Evaluation." *American Journal of Evaluation* 9(3):5-14.

Heider, C. 2011. "A Conceptual Framework for Developing Evaluation Capacities: Building on Good Practice." in *Influencing Change. Building Evaluation Capacity to Strengthen Governance*, edited by R. C. Rist, M. H. Boily, and F. Martin. Washington, D.C., US: The World Bank.

Heilman, J. G. 1983. "Beyond the Technical and Bureaucratic Theories of Utilization: Some Thoughts on Synthesizing Reviews and the Knowledge Base of the Evaluation Profession " *Evaluation Review* 7(6):707-28.

Henkel, M. 1991. *Government, Evaluation and Change*. London, UK: Jessica Kingsley Publishers.

Henry, G. T. 2001. "How Modern Democracies Are Shaping Evaluation and the Emerging Challenges for Evaluation." *American Journal of Evaluation* 22(3):419-29.

Henry, G. T., and M. M. Mark. 2003a. "Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions." *American Journal of Evaluation* 24(3):293-314.

—. 2003c. "Toward an Agenda for Research on Evaluation." *New Directions for Evaluation* 2003(97):69-80.

Héritier, A., and D. Lehmkuhl. 2008. "The Shadow of Hierarchy and New Modes of Governance." *Journal of Public Policy* 28(1):1-17.

Hernández Licona, G. 2010. "Construyendo un Sistema de Evaluación y Medición: Un Reto de Política Pública con Elementos Técnicos."

Hertting, N., and E. Vedung. 2012. "Purposes and criteria in network governance evaluation: How far does standard evaluation vocabulary takes us?" *Evaluation* 18(1):27-46.

Hill, M. J. (Ed.). 1997. *The Policy Process: A reader*. London, UK: Routledge.

—. 2005. *The Public Policy Process*. London, UK: Routledge.

Hitchcock, G. 2011. "Care Quality Commission faces uphill struggle, says watchdog." in *The Guardian*. London, UK.

HM Government. 2010. "The Coalition: Our programme for government." London, UK.

—. 2012. "Open Data White Paper. Unleashing the potential.", edited by Cabinet Office. London, UK.

HM Treasury. 2010. "Spending Review 2010." edited by HM Treasury. London, UK.

—. 2011a. "The Green Book. Appraisal and Evaluation in Central Government." London, UK.

—. 2011c. "The Magenta Book. Guidance for evaluation." edited by HM Treasury. London, UK.

—. 2012a. "Treasury Minutes. Progress on implementing recommendations from 28 Committee of Public Accounts reports (Session 2010-12) and 5 National Audit Office reports." London, UK.

—. 2012b. "Treasuy Minutes. Government responses on the Fifty Sixth, the Seventy First, the Seventy Sixth and the Seventy Eighth Reports from the Committee of Public Accounts: Session 2010-12.". London, UK.

—. 2013. "Spending Round 2013." London, UK.

Hodkinson, P., and H. Hodkinson. 2001. "The Strengths and Limitations of Case Study Research." in *Making an Impact on Policy and Practice*. Cambridge, UK.

Hoefer, R. 2000. "Accountability in Action?: Program Evaluation in Nonprofit Human Service Agencies." *Nonprofit Management and Leadership* 11(2):167-77.

—. 2012. *Policy Creation and Evaluation: understanding welfare reform in the United States*. United States: Oxford University Press.

Hogwood, B. W., and L. A. Gunn. 1984. *Policy Analysis for the Real World*. Oxford, UK: Oxford University Press.

Højlund, S. 2014. "Evaluation use in the organizational context – changing focus to improve theory." *Evaluation* 20(1):26-43.

Holliday, I. 2000. "Is the British State Hollowing Out?" *The Political Quarterly* 71(2):167-76.

Hood, C. 1991. "A public management for all seasons?" *Public Administration* 69:3-19.

—. 2002. "The Risk Game and the Blame Game." *Government and Opposition* 37(1):15-37.

—. 2011. *The Blame Game: Spin, bureaucracy, and self-preservation in government*. New Jersey, US: Princeton University Press.

—. 2014. "Accountability and Blame-Avoidance." in *The Oxford Handbook of Public Accountability*, edited by M. Bovens, R. E. Goodin, and T. Schillemans. Oxford, UK.

Horton, S. 2006. "The Public Service Ethos in the British Civil Service: An Historical Institutional Analysis." *Public Policy and Administration* 21(1):32-48.

House, E. R. 1980. *Evaluating with Validity*. London, UK: SAGE.

—. 1997. "Evaluation in the Government Marketplace." *American Journal of Evaluation* 18(1):37-48.

—. 2006. "Democracy and Evaluation." *Evaluation* 12(1):119-27.

—. 2008. "Blowback: Consequences of Evaluation for Evaluation." *American Journal of Evaluation* 29(4):416-26.

House, E. R., and K. R. Howe. 1998. "The Issue of Advocacy in Evaluations." *The American Journal of Evaluation* 19(2):233-36.

—. 1999. *Values in Evaluation and Social Research*. London, UK: Sage.

Howell, E. M., and A. Yemane. 2006. "An Assessment of Evaluation Designs: Case Studies of 12 Large Federal Evaluations." *American Journal of Evaluation* 27(2):219-36.

Howlett, B. W., and A. G. Lewis. 1984. *Policy Analysis for the Real World*. Oxford, UK: Oxford University Press.

Howlett, M. 2000. "Managing the "hollow state": procedural policy instruments and modern governance." *Canadian Public Administration-Administration Publique Du Canada* 43(4):412-31.

Huberman, A. M., and M. B. Miles. 2002. *The qualitative researcher's companion*. London, UK: SAGE.

Hughes, M., and T. Traynor. 2000. "Reconciling Process and Outcome in Evaluating Community Initiatives." *Evaluation* 6(1):37-49.

Hunter, D. E. K., and S. B. Nielsen. 2013. "Performance Management and Evaluation: Exploring Complementarities." *New Directions for Evaluation* 2013(137):7-17.

Independent Commission for Aid Impact. 2015. "ICAI." London, UK.

Instituto de Innovación Educativa. 2013. "Evaluación Externa de los Resultados del Consejo Nacional de Evaluación de la Política de Desarrollo Social." Mexico City: Instituto Tecnológico y de Estudios Superiores de Monterrey.

Inter-American Development Bank. 2008. "Results-based Budget Model Support Programme-Phase I." Washington, D.C., US.

International Federation of Accountants. 2010. "Handbook of International Quality Control, Auditing, Review, other Assurance, and Related Services Pronouncements." New York, USA.

Ipsos Mori. 2015. "About us." London, UK.

Iriti, J. E., W. E. Bickel, and C. Awsumb Nelson. 2005. "Using Recommendations in Evaluation." *American Journal of Evaluation* 26(4):464-79.

ITV News. 2015. "British Medical Association chief warns next government against introducing charges on the NHS." London, UK.

Jacob, S. 2011. "Sharing Power among Evaluation Players: Mission Possible?" in *Evaluation. Seeking Truth or Power?*, edited by P. Eliadis, J. E. Furubo, and S. Jacob. New Jersey, US: Transaction Publishers.

Jacob, S., S. Speer, and J. E. Furubo. 2015. "The institutionalization of evaluation matters: Updating the International Atlas of Evaluation 10 years later." *Evaluation* 21(1):6-31.

Jaime, E. 2014. "Goodbye CONEVAL?". Mexico City.

James, O. 2004. "The UK Core Executive's Use of Public Service Agreements as a Tool of Governance." *Public Administration* 82(2):397-419.

Jann, W., and K. Wegrich. 2007. "Theories of the Policy Cycle." in *Handbook of Public Policy Analysis. Theory, Politics and Methods*, edited by F. Fisher, G. J. Miller, and M. S. Sidney. US: CRC Press.

Jick, T. D. 1979. "Mixing Qualitative and Quantitative Methods: Triangulation in Action." *Administrative Science Quarterly* 24(4):602-11.

John, P. 1999. *Analysing Public Policy*. London, UK: Continuum.

Johnson, C., and C. Talbot. 2007. "The UK Parliament and performance: challenging or challenged?" *International Review of Administrative Sciences* 73(1):113-31.

Johnson, K., L. O. Greenseid, S. A. Toal, J. A. King, F. Lawrenz, and B. Volkov. 2009. "Research on Evaluation Use. A Review of the Empirical Literature From 1986 to 2005." *American Journal of Evaluation* 30(3):377-410.

Johnson, R. B., and A. J. Onwuegbuzie. 2004. "Mixed Methods Research: A Research Paradigm Whose Time Has Come." *Educational Researcher* 33(7):14-26.

Jones, H. 2012. "Promoting evidence-based decision-making in development agencies." Overseas Development Institute.

Jordan, A. J., and J. R. Turnpenny (Eds.). 2015. *The Tools of Policy Formulation: Actors, Capacities, Venues and Effects*. UK: Edward Elgar Publishing.

Juárez-Ramírez, C., M. Márquez-Serrano, N. Salgado de Snyder, B. E. Pelcastre-Villafuerte, M. G. Ruelas-González, and H. Reyes-Morales. 2014. "La desigualdad en salud de grupos vulnerables de México: Adultos mayores, indígenas y migrantes." *Revista Panamericana de Salud Pública* 35:284-90.

Jupp, V. 2006. *The SAGE Dictionary of Social Research Methods*. London, UK: SAGE.

Kaarbo, J., and R. K. Beasley. 1999. "A Practical Guide to the Comparative Case Study Method in Political Psychology." *Political Psychology* 20(2):369-91.

Karlsson, O. 1996. "A Critical Dialogue in Evaluation: How Can the Interaction between Evaluation and Politics be Tackled?" *Evaluation* 2(4):405-16.

—. 2003. "Evaluation Politics in Europe: Trends and Tendencies." *Studies in Educational Policy and Educational Philosophy*.

Keen, J., and T. Packwood. 1995. "Case Study Evaluation." *BMJ: British Medical Journal* 311(7002):444-46.

Kennett, P. (Ed.). 2008. *Governance, Globalization and Public Policy*. Cheltenham, UK: Edward Elgar Publishing Ltd.

Kernaghan, K. 2000. "The Post-Bureaucratic Organization and Public Service Values." *International Review of Administrative Sciences* 66(1):91-104.

Kinder, D. R., and T. R. Palfrey. 1993. *Experimental Foundations of Political Science*: University of Michigan Press.

King, G., R. O. Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. New Jersey, US: Princeton University Press.

Kingdon, J. W. 1995. *Agendas, Alternatives and Public Policies*. New York, US: Harper Collins.

Kniker, T. 2011. "Evaluation Survivor: How to Outwit, Outplay, and Outlast as an Internal Government Evaluator." *New Directions for Evaluation* 2011(132):57-72.

Knill, C., and J. Tosun. 2012. *Public Policy. A New Introduction*. New York, US: Palgrave Macmillan.

Knoepfel, P., C. Larrue, F. Varone, and M. Hill. 2011. *Public Policy Analysis*. London, UK: Policy Press.

Krathwohl, D. R. 1980. "The Myth of Value-Free Evaluation." *Educational Evaluation and Policy Analysis* 2(1):37-45.

Kudo, H. 2003. "Between the 'Governance' Model and the Policy Evaluation Act: New Public Management in Japan." *International Review of Administrative Sciences* 69(4):483-504.

Kuzemko, C. 2014. "Politicising UK energy: what 'speaking energy security' can do." *Policy & Politics* 42(2):259-74.

Labour Party. 2014. "Patients before profits."

Labour Press. 2015. "Response to PAC report on deteriorating cancer services."

Lasswell, H. D. 1970. "The emerging conception of the policy sciences." *Policy Sciences* 1(1):3-14.

Lawrence Neuman, W. 2014. *Basics of Social Research: Qualitative & Quantitative Approaches*. London, UK: Pearson Education Limited.

Lawrenz, F., A. Gullickson, and S. Toal. 2007. "Dissemination. Handmaiden to Evaluation Use." *American Journal of Evaluation* 28(3):275-89.

Laws, D., and M. Hajer. 2006. "Policy in Practice." in *The Oxford Handbook of Public Policy*, edited by M. Moran, M. Rein, and R. E. Goodin. Oxford, UK: Oxford University Press.

Ledermann, S. 2011. "Exploring the Necessary Conditions for Evaluation Use in Program Change." *American Journal of Evaluation*.

Leeuw, F. L. 1991. "Policy theories, knowledge utilization, and evaluation." *Knowledge and Policy* 4(3):73-91.

—. 1996. "Auditing and Evaluation: Bridging a Gap, Worlds to Meet?" *New Directions for Evaluation* 1996(71):51-60.

—. 2009. "Evaluation policy in the Netherlands." *New Directions for Evaluation* 2009(123):87-102.

Leeuw, F. L., J. Toulemonde, and A. Brouwers. 1999. "Evaluation Activities in Europe: A Quick Scan of the Market in 1998." *Evaluation* 5(4):487-96.

Lennie, J., J. Tacchi, M. Wilmore, and B. Koirala. 2015. "A holistic, learning-centred approach to building evaluation capacity in development organizations." *Evaluation* 21(3):325-43.

Leviton, L. C., and E. F. X. Hughes. 1981. "Research On the Utilization of Evaluations: A Review and Synthesis." *Evaluation Review* 5(4):525-48.

Levy, R. P., M. Barzelay, and A. M. Porras-Gomez. 2011. "The Reform of Financial Management in the European Commission: A Public Management Policy Cycle Case Study." *Public Administration* 89(4):1546-67.

Liberal Democrat Party. 2014. "Liberal Democrat Pre-Manifesto 2014."

Lindblom, C. E. 1959. "The Science of "Muddling Through"." *Public Administration Review* 19(2):79-88.

Ling, T. 2002. "Delivering Joined-Up Government in the UK: Dimensions, Issues and Problems." *Public Administration* 80(4):615-42.

Lion, C., P. Martini, and S. Volpi. 2006. "Evaluating the Implementation Process: A Contribution within the Framework of the European Social Fund (ESF) Programme." *Evaluation* 12(3):313-29.

Lipsky, D. B., R. L. Seeber, A. C. Avgar, and R. M. Scanza. 2007. "Managing the Politics of Evaluation: Lessons from the Evaluation of ADR Programs." Pp. 116-29 in *Proceedings of the Fift-Ninth Annual Meeting of the Labor and Employment Relations Association*.

Lipsky, M. 1971. "Street-Level Bureaucracy and the Analysis of Urban Reform." *Urban Affairs Review* 6(4):391-409.

Lloyd, N., and L. Harrington. 2012. "The challenges to effective outcome evaluation of a national, multi-agency initiative: The experience of Sure Start." *Evaluation* 18(1):93-109.

Lobel, O. 2012. "New Governance as Regulatory Governance." in *The Oxford Handbook of Governance*, edited by D. Levi-Faur. Oxford, UK: Oxford University Press.

London School of Economics. 2014. "GV314 Empirical Research in Government."

Long, R. 2015. "Finding a Path for the Cure for Dementia. An independent report into an integrated approach to dementia research.". London, UK: Department of Health.

Lonsdale, J. 2008. "Balancing Independence and Responsiveness: A Practitioner Perspective on the Relationships Shaping Performance Audit." *Evaluation* 14(2):227-48.

Lopez-Acevedo, G., P. Krause, and K. Mackay (Eds.). 2012. *Building Better Policies. The Nuts and Bolts of Monitoring and Evaluation Systems.* Washington, D.C., US: The World Bank.

Lozano, E., N. Gutiérrez, R. Díaz, and A. González. 2007. "Adquisición de bienes privados con recursos públicos. Revisión de Política Agropecuaria y Rural en México.": Banco Interamericano de Desarrollo.

Luo, H. 2010. "The Role for an Evaluator: A Fundamental Issue for Evaluation of Education and Social Programs " *International Education Studies* 3(2):42-50.

Lustig, N. 2011. "Scholars who became practitioners. The influence of research on the design, evaluation and political suvival of Mexico's Antipoverty Program Progresa/Oportunidades." edited by Center for Global Development.

Mackay, K. 2007. "How to build M & E systems to support better government." Pp. ix, 157 p. Washington, D.C., US: The World Bank.

Mactaggart, F. MP. 2014. "Slough MP 'Wears it Pink' for Breast Cancer Awareness."

Mahoney, J. 2007. "Qualitative Methodology and Comparative Politics." *Comparative Political Studies* 40(2):122-44.

Mahoney, J., and G. Goertz. 2006. "A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research." *Political Analysis* 14(3):227-49.

Maldonado, C. 2011. "El programa SICALIDAD: Ejemplo de gestión pública anclada en la evaluación para la mejora continua." in *De las recomendaciones a las acciones: La experiencia del Premio 2011. Programas Federales Comprometidos con el Proceso de Evaluación*, edited by G. Cejudo and C. Maldonado. Mexico City: CIDE-SFP.

—. 2012. "Más allá de las instituciones: La evaluación y la gestión por resultados en México. Una perspectiva de mercado." in *XVII Congreso Internacional del CLAD sobre la Reforma del Estado y de la Administración Pública*. Cartagena, Colombia.

Malena, C., and M. McNeil. 2010. "Social Accountability in Africa: An Introduction." in *Demanding Good Governance: Lessons from Social Accountability Initiatives in Africa*, edited by C. Malena and M. McNeil. Washington, D.C., US: The World Bank.

Mark, M. M. 2001. "Evaluation's future: furor, futile, or fertile?" *The American Journal of Evaluation* 22(3):457-79.

Mark, M. M., L. J. Cooksy, and W. M. K. Trochim. 2009. "Evaluation policy: An introduction and overview." *New Directions for Evaluation* 2009(123):3-11.

Marston, G., and R. Watts. 2003. "Tampering With the Evidence: A Critical Appraisal of Evidence-Based Policy-Making." *The Drawing Board: An Australian Review of Public Affairs* 3(3):143-63.

Martín del Campo Rodríguez, A. C. 2010-2011. "Specific Performance Evaluation. Health Caravans Programme.". Mexico City.

Martin, S., J. Downe, C. Grace, and S. Nutley. 2010. "Validity, Utilization and Evidence-Based Policy: The Development and Impact of Performance Improvement Regimes in Local Public Services." *Evaluation* 16(1):31-42.

Martínez, N., and A. Rivera. 2015. "Less autonomy for CONEVAL." in *El Universal*. Mexico City.

Marvick, D. 1980. "The Work of Harold D. Lasswell: His Approach, Concerns, and Influence." *Political Behavior* 2(3):219-29.

Mason, J. 2002. *Qualitative Researching*. London, UK: Sage.

Mason, R. 2015. "Brief, brutal and very public: there's more to Margaret Hodge's grillings than dramatics." in *The Guardian*. London, UK.

Mathison, S. 2000. "Deliberation, Evaluation, and Democracy." *New Directions for Evaluation* (85):85-89.

Matthews, F. 2013. *Complexity, Fragmentation, and Uncertainty. Government Capacity in an Evolving State.* Oxford, UK: Oxford University Press.

May, T. 2011. "Social Research: Issues, Methods and Research." California, US: Open University Press.

Mazmanian, D. A., and P. A. Sabatier. 1983. *Implementation and Public Policy*. Illinois, US: University Press of America.

McAnulla, S. 2006. *British Politics: A Critical Introduction*. London, UK: Continuum.

McDavid, J. C., and I. Huse. 2012. "Legislator Uses of Public Performance Reports." *American Journal of Evaluation* 33(1):7-25.

McDonald, B., P. Rogers, and B. Kefford. 2003. "Teaching People to Fish? Building the Evaluation Capability of Public Sector Organizations." *Evaluation* 9(1):9-29.

McKie, R. 2015. "David Cameron's flagship Cancer Drugs Fund 'is a waste of NHS cash' " in *The Guardian*. London, UK.

McKinnell, C., MP,. "Catherine goes pink for breast cancer awareness."

McLean, I., D. Haubrich, and R. Gutiérrez-Romero. 2007. "The Perils and Pitfalls of Performance Measurement: The CPA Regime for Local Authorities in England." *Public Money & Management* 27(2):111-18.

McLemore, J. R., and J. E. Neumann. 1987. "The inherently political nature of program evaluators and evaluation research." *Evaluation and Program Planning* 10(1):83-93.

McMillan Cancer Support. "Lack of social care causing devastating consequences for 100000s of cancer patients."

—. 2015. "Macmillan Cancer Support responds to NAO report on cancer services." London, UK.

McMillan Cancer Support, Department of Health, and Cancer Research UK. 2012. "Cancer Services Coming of Age: Learning from the Improving Cancer Treatment Assessment and Support for Older People Project." London, UK.

McSmith, A. 2015. "Andy McSmith's Diary: So that's how Meg Hillier became Margaret Hodge's successor." in *The Independent*. London, UK.

Melkers, J., and D. Roessner. 1997. "Politics and the Political Setting as an Influence on Evaluation Activities: National Research and Technology Policy Programs in the United States and Canada." *Evaluation and Program Planning* 20(1):57-75.

Mensa-Bonsu, I. F., and M. Holm Andersen. 2010. "Involving Developing Countries in the Evaluation of Development Cooperation: The Case of a Joint Evaluation of Development Cooperation between Ghana and Denmark." *Evaluation* 16(3):295-308.

Mény, Y., and J. C. Thoening. 1992. *Las políticas públicas*. Barcelona, España: Ariel.

Merino, M. 2009. "Informe sobre la evolución y el desempeño de la Auditoría Superior de la Federación." Mexico City.

Meyer, C. B. 2001. "A Case in Case Study Methodology." *Field Methods* 13(4):329-52.

Miller, T., M. Birch, M. Mauthner, and J. Jessop (Eds.). 2012. *Ethics in Qualitative Research*. London, UK: SAGE.

Millett, D. 2015. "CQC's 9% fee hike is a 'financial blow' to GP practices." London, UK.

Mills, A. J., G. Durepos, and E. Wiebe. 2010. "Case Study as a Methodological Approach. Encyclopedia of Case Study Research." Pp. 67-78. London, UK: SAGE.

Milstein, B., T. J. Chapel, S. F. Wetterhall, and D. A. Cotton. 2002. "Building Capacity for Program Evaluation at the Centers for Disease Control and Prevention." *New Directions for Evaluation* 2002(93):27-46.

Ministry of Finance. "Cédula: Conceptos Básicos del Presupuesto Basado en Resultados. Sistema de Evaluación del Desempeño (PbR-SED)."

—. 2011. "Results-based Budget. Evaluation.".

Ministry of Health. 2005. "Sistema de Protección Social en Salud. Estrategia de Evaluación (Evaluation Strategy)." Mexico City.

—. 2007a. "Sectorial Health Programme 2007-2012 ", edited by Ministry of Health. Mexico City.

—. 2007b. "Specific Action Programme 2007-2012 SICALIDAD." Mexico City.

—. 2011a. "Health Caravans Programme. Specific Action Plan 2007-2012." Mexico, City.

—. 2011b. "Specific Performance Evaluation 2010-2011. Health Caravans Programme. Institutional Statement.". Mexico City.

—. 2011-2012. "Consistency and Results Evaluation. Health Caravans Programme.". Mexico City.

—. 2013a. "Agreement that issues the operational rules of the Integral Health Quality System Programme for the fiscal year 2014." Mexico City.

—. 2013b. "Health Caravans Evaluations."

—. 2014a. "Follow-up mechanism of recommendations. Work plan. Health Caravans Programme.".

—. 2014b. "Sistema Integral de Calidad en Salud SICALIDAD. Work Programme." Mexico City.

Ministry of Public Function. 2014. "Government Performance Evaluation Unit." Mexico City.

Mintzberg, H. 1985. "The organization as political arena." *Journal of Management Studies* 22(2):133-54.

Miyakawa, T. 1999. *The Science of Public Policy*. London, UK: Routledge.

Mohan, R. 2014. "Evaluator Advocacy: It Is All in a Day's Work." *American Journal of Evaluation* 35(3):1-7.

Mohan, R., and K. Sullivan. 2006. "Managing the Politics of Evaluation to Achieve Impact." *New Directions for Evaluation* 2006(112):7-23.

Molas-Gallart, J. 2012. "Research Governance and the Role of Evaluation: A Comparative Study." *American Journal of Evaluation* 33(4):583-98.

Monaghan, M. 2011. *Evidence versus Politics: Exploiting Research in UK Drug Policy Making?* Bristol, UK: Bristol Policy.

Moody, J. 2015. "Breast screening unit is now back where it belongs in Uttoxeter." in *Uttoxeter Adviser*.

Moran, M. 2003. *The British Regulatory State: High Modernism and Hyper-innovation*. Oxford, UK: Oxford University Press.

Morra, L. G., and A. C. Friedlander. 1999. "Case Study Evaluations." edited by The World Bank.

Morris, M. 1990. "A Nightmare in Elm City: When Evaluation Field Experiences Meet Organizational Politics." *Evaluation Review* 14(1):91-99.

Morris, S. 2005. "Evaluating Public Policy in the UK: History, Politics & Practice."

Morton, M. H. 2009. "Applicability of Impact Evaluation to Cohesion Policy." in *An Agenda for a Reformed Cohesion Policy*, edited by European Commission.

Moses, J. W., and T. L. Knutsen. 2012. *Ways of Knowing: Competing Methodologies in Social and Political Research*. New York, US: Palgrave Macmillan.

Mosse, R., and L. E. Sontheimer. 1996. "Performance Monitoring Indicators Handbook." Washington, D.C., US: The World Bank.

Moynihan, D. P. 2006. "Managing for Results in State Government: Evaluating a Decade of Reform." *Public Administration Review* 66(1):77-89.

Mulgan, G., and A. Lee. 2001. "Better Policy Delivery and Design: A Discussion Paper." edited by Performance and Innovation Unit. London, UK.

My M&E. 2004. "Purposes of Evaluation."

Nagel, S. (Ed.). 2002. *The Handbook of Public Policy Evaluation*. California, US: SAGE.

NatCen Social Research. "British Social Attitudes Survey 32." London, UK.

National Audit Office. "Major Outputs." London, UK.

—. "Planning the value for money programme." London, UK.

—. "Quality and Independence." London, UK.

—. "What is a Value for Money Study?". London, UK.

—. 2010. "Delivering the Cancer Reform Strategy." London, UK.

—. 2011a. "The Care Quality Commission: Regulating the quality and safety of health and adult social care." London, UK.

—. 2011e. "The Care Quality Commission: Regulating the quality and safety of health and adult social care. Detailed Methodology.". London, UK: National Audit Office.

—. 2012. "The management of adult diabetes services in the NHS." London, UK.

—. 2013a. "Evaluation in government." edited by National Audit Office. London, UK.

—. 2013j. "History of the NAO." London, UK.

—. 2013k. "An Introduction to the National Audit Office." London, UK.

—. 2013n. "National Audit Office. Annual Report and Accounts 2012-13." edited by National Audit Office. London, UK.

—. 2014a. "National Audit Office. Annual Report and Accounts 2013 -14." edited by National Audit Office. London, UK.

—. 2014c. "The Privatisation of Royal Mail." London, UK: National Audit Office.

—. 2015a. "Capacity and capability to regulate the quality and safety of health and adult social care." London , UK.

—. 2015b. "Progress in improving cancer services and outcomes in England." London, UK.

—. 2015d. "Whole of Government Accounts 2013-14." London, UK.

National Cancer Intelligence Network. 2012. "Recurrent and Metastatic Breast Cancer Data Collection Project. Pilot Report.".

National Council for the Evaluation of the Social Development Policy. 2011a. "II Award "Good practices in the use of monitoring and evaluation outputs in policymaking" 2010-2011."

—. 2011b. "Terms of Reference Model for the Consistency and Results Evaluation." Mexico City.

—. 2012a. "About us." Mexico City.

—. 2012b. "Dissemination Index of External Evaluations to Social Development Federal Programmes."

—. 2014a. "Coneval analyses the access and effective use of health services in population affiliated to *Seguro Popular* at the national level 2012." Mexico City.

—. 2014b. "Considerations for the 2015 Budget Process." Mexico City.

—. 2014c. "Factual and legal autonomy: The challenges imposed by the political reform to CONEVAL." Mexico City.

—. 2014d. "Monitoring Record Card."

—. 2015. "Terms of Reference Model for the Design Evaluation." Mexico City.

National Health Service. 2012. "The NHS Outcomes Framework 2013/14." edited by National Health Service. London, UK.

—. 2014. "Improving outcomes: a strategy for cancer." London, UK.

National Institute for Health and Care Excellence. 2014. "Who we are." London, UK.

National Institute for Health Research. 2015. "Health Technology Assessment (HTA) Programme." London, UK.

National Public Health Institute. 2012. "Evaluation of the effects of the catastrophic expenses fund in breast cancer patients ". Mexico City.

—. 2013a. "Specific Performance Evaluation 2012-2013 SICALIDAD." Mexico City.

—. 2013b. "Specific Performance Evaluation. Health Caravans Programme 2012-2013.". Mexico City.

—. 2013f. "Specific Performance Evaluation. SICALIDAD Programme.". Mexico City.

Nelson, B. J. 1984. *Making an Issue of Child Abuse: Political Agenda Setting for Social Problems*. Illinois, US: University of Chicago Press.

Nelson, L. 2008. "Utilization-focused Evaluation of the Program Evaluation Process of a Missouri School District." Missouri, US: University of Missouri.

Neuman, A., N. Shahor, I. Shina, A. Sarid, and Z. Saar. 2013. "Evaluation utilization research—Developing a theory and putting it to use." *Evaluation and Program Planning* 36(1):64-70.

Newcomer, K. E., and S. L. Caudle. 1991. "Evaluating Public Sector Information Systems: More Than Meets the Eye." *Public Administration Review* 51(5):377-84.

Noor, K.B.M. 2008. "Case Study: A Strategic Research Methodology." *American Journal of Applied Sciences* 5(11):1602-04.

Norris, N. 2005. "The Politics of Evaluation and the Methodological Imagination." *American Journal of Evaluation* 26(4):584-86.

—. 2007. "Evaluation and Trust." Pp. 139-53 in *Dilemmas of Engagement: Evaluation and the New Public Management*, edited by S. Kushner and N. Norris: JAI Press Inc.

Nutley, S., H. Davies, and I. Walter. 2002. "Evidence Based Policy and Practice: Cross Sector Lessons From the UK." in *Social Policy Research and Evaluation Conference*. Wellington, New Zealand.

O'Brien, T., S. Payne, M. Nolan, and C. Ingleton. 2010. "Unpacking the Politics of Evaluation: A Dramaturgical Analysis." *Evaluation* 16(4):431-44.

O'Brien, Z. 2015. "Whipps Cross Hospital has been placed in special measures by the Care Quality Commission over bullying." in *Guardian*.

O'Hara, M., and J. Shepherd. 2011. "Government cuts millions from research and evaluation budgets " in *The Guardian*. London, UK.

O'Leary, Z. 2004. *The Essential Guide to Doing Research*. London, UK: SAGE.

O'Toole, L. J. Jr. 2000. "Research on Policy Implementation: Assessment and Prospects." *Journal of Public Administration Research and Theory* 10(2):263-88.

OECD. 2002. "Relations Between Supreme Audit Institutions and Parliamentary Committees." in *Sigma Papers*.

—. 2005. *Modernising Government: The Way Forward*. Paris, France: OECD.

Oliver, K., T. Lorenc, and S. Innvær. 2014. "New directions in evidence-based policy research: A critical analysis of the literature." *Health Research Policy and Systems* 12(1).

Orr, S. K. 2010. "Exploring Stakeholder Values and Interests in Evaluation." *American Journal of Evaluation* 31(4):557-69.

Osborne, D., and P. Plastrik. 1998. *Banishing Bureaucracy: The Five Strategies for Reinventing Government*. New York, US: Plume.

Ovarian Cancer Action. 2013. "Annual Report and Audited Accounts for the Year Ending 31st March." London, UK: Ovarian Cancer Action.

Owen, J., T. Cook, and E. Jones. 2005. "Evaluating the Early Excellence Initiative: The Relationship between Evaluation, Performance Management and Practitioner Participation." *Evaluation* 11(3):331-49.

Owen, J. M., and P. J. Rogers. 1999. *Program Evaluation. Forms and approaches*. Melbourne, Australia: SAGE.

Painter, C. 2012. "The UK Coalition government: Constructing public service reform narratives." *Public Policy and Administration* 20(1):3-20.

Palumbo, D. J. 1987. *The Politics of Program Evaluation*. California, US: SAGE.

Palumbo, D. J., and M. A. Harder (Eds.). 1981. *Implementing Public Policy*. Massachusetts, US: Lexington Books.

Panchamia, N., and P. Thomas. 2014. "Public Service Agreements and the Prime Minister's Delivery Unit." Institute for Government.

Parliament UK. "Committee inquiries: Open calls for evidence." London, UK: House of Commons.

—. "Guidance on giving evidence to a Select Committee of the House of Commons ". London, UK: House of Commons.

—. "What are Early day motions? ." London, UK.

—. 2013. "Early Day Motion 1124 Cancer Research UK." London, UK.

Parsons, W. 1995. *Public Policy: An Introduction to the Theory and Practice of Policy Analysis*. London, UK: Edward Elgar Publishing Ltd.

—. 2002. "From Muddling Through to Muddling Up - Evidence Based Policy Making and the Modernisation of British Government." *Public Policy and Administration* 17(3):43-60.

—. 2007. *Políticas Públicas*. Mexico City: Flacso México.

Partridge, E. 2011. "Caught in the Same Frame? The Language of Evidence-based Policy in Debates about the Australian Government 'Intervention' into Northern Territory Aboriginal Communities." in *Evidence And Evaluation in Social Policy*, edited by I. Greener and B. Greve. London, UK: Wiley Blackwell.

Patten, M. L. 2007. *Understanding Research Methods: An Overview of the Essentials*. California, US: Pyrczak Publishing.

Patton, M. Q. 1986. *Utilization-focused evaluation*. London, UK: SAGE.

—. 1987. "Evaluation's Political Inherency: Practical Implications for Design and Use." in *The Politics of Program Evaluation*, edited by Dennis J. Palumbo. California, US: SAGE.

Pawson, R., and N. Tilley. 1997. *Realistic Evaluation*. London, UK: SAGE.

Pearce, W., and S. Raman. 2014. "The new randomised controlled trials (RCT) movement in public policy: challenges of epistemic governance." *Policy Sciences* 47(4):387-402.

Pelaez, M. 2015. "Aureoles pide que la ASF tenga dientes para sancionar (Aureoles asks for "teeth" for the SAO to sanction)." Azteca Noticias.

Pérez-Jácome Friscione, D. 2012. "Presupuesto basado en resultados: Origen y aplicación en México." edited by Ministry of Finance. Mexico City.

Pérez, L. 2015. "Coahuila spent only 36% of the Health Caravans funding." in *El Siglo*. Coahuila, Mexico.

Pérouse de Montclos, M. A. 2012. "Humanitarian action in developing countries: Who evaluates who?" *Evaluation and Program Planning* 35(1):154-60.

Perri 6, and C. Bellamy. 2012. *Principles of Methodology: Research Design in Social Science*. Los Angeles: SAGE.

Peters, B. G. 2001. *The Politics of Bureaucracy*. London, UK: Routledge.

—. 2012. "Information and Governing: Cybernetic Models of Governance." in *The Oxford Handbook of Governance*, edited by D. Levi-Faur. Oxford, UK: Oxford University Press.

—. 2015. *Advanced Introduction to Public Policy*. Cheltenham, UK: Edward Elgar Publishing.

Peters, B. G., and J. Pierre. 2012. *The SAGE Handbook of Public Administration*. Los Angeles, US: SAGE.

Pharmaphorum. 2015. "Pharma 'shocked' as drugs cut from Cancer Drugs Fund."

Picciotto, R. 1999. "Towards an Economics of Evaluation." *Evaluation* 5(1):7-22.

—. 2005. "Policy Coherence and Development Evaluation: Issues and Possible Approaches." in *Fostering Development in a Global Economy: A Whole of Government Perspective*, edited by OECD.

—. 2011. "The logic of evaluation professionalism." *Evaluation* 17(2):165-80.

—. 2013. "Evaluation Independence in Organizations." *Journal of MultiDisciplinary Evaluation* 9(20):18-32.

—. 2015. "Democratic evaluation for the 21st century." *Evaluation* 21(2):150-66.

Pierre, J. 2012. "Governance and Institutional Flexibility." in *The Oxford Handbook of Governance*, edited by D. Levi-Faur. Oxford, UK: Oxford University Press.

Pierre, J., and B. G. Peters. 2000. *Governance, Politics and the State*. Basingstoke, UK: Palgrave Macmillan.

Pierson, P. 2007. "The Costs of Marginalization: Qualitative Methods in the Study of American Politics." *Comparative Political Studies* 40(2):146-69.

Pollitt, C. 1993. "Occasional Excursions: A Brief History of Policy Evaluation in the UK." *Parliamentary Affairs* 46(3):353-62.

—. 1998. "Evaluation in Europe: Boom or Bubble?" *Evaluation* 4(2):214-24.

—. 1999. "The Changing Context." in *Performance or compliance?: Performance Audit and Public Management in Five Countries*, edited by C. Pollitt, X. Girre, J. Londsdale, R. Mul, H. Summa, and M. Waerness. Oxford, UK: Oxford University Press.

—. 2006. "Performance Information for Democracy: *The Missing Link*?" *Evaluation* 12(1):38-55.

—. 2013. "The logics of performance management." *Evaluation* 19(4):346-63.

Pollitt, C., and G. Bouckaert. 2000. *Public Management Reform: A Comparative Analysis*. Oxford, UK: Oxford University Press.

Pollitt, C., and H. Summa. 1996. "Performance Audit and Evaluation: Similar Tools, Different Relationships?" *New Directions for Evaluation* 1996(71):29-50.

Pollitt, C., and C. Talbot. 2004. *Unbundled Government: A critical analysis of the global trend to agencies, quangos and contractualisation*. London, UK: Routledge.

Posner, P. L., and R. Schwarzt. 2011. "Accountability Institutions and Information in the Policy-Making Process." in *Accountable Governance: Problems and Promises*, edited by M. J. Dubnick and H. G. Frederickson. London, UK: Routledge.

Post & Parcel. 2014. "NAO says UK lost hundreds of millions through Royal Mail IPO." London, UK.

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

Presidencia de la República. "Plan Nacional de Desarrollo 2013-2018." Mexico City.

Press Associated. 2013. "Care Quality Commission 'not fit for purpose' says BMA." in *The Guardian*. London, UK.

Pressman, J. L., and A. B. Wildavsky. 1984. *Implementation: How great expectations in Washington are dashed in Oakland or, why it's amazing that federal programs work at all, this being a saga of the Economic Development Administration as told by two sympathetic observers who seek to build morals on a foundation of ruined hopes*. California, US: University of California Press.

Prime Minister and the Minister for the Cabinet Office. 1999. "Modernising Government."

Prince, M. J. 2000. "Banishing Bureaucracy or Hatching a Hybrid? The CanadianFood Inspection Agency and the Politics of Reinventing Government." *Governance* 13(2):215-32.

Proctor, R. 1995. *Cancer wars: How Politics Shapes What We Know And Don't Know About Cancer*. New York, US.: BasicBooks.

Propper, C., and D. Wilson. 2003. "The Use and Usefulness of Performance Measures in the Public Sector." *CMPO Working Paper Series* 03(073).

Public Administration Select Committee. 2003. "On Target? Government By Measurement." London, UK.

Public Health England. 2013. "Standard Evaluation Frameworks." London, UK.

—. 2014a. "About us." London, UK.

—. 2014b. "Cold weather plan for England 2014." London, UK.

Punch, K. F. 2014. *Introduction to Social Research: Quantitative and Qualitative Approaches*: California, US.

PWC. 2012. "The Business of Evidence. A report prepared for the Market Research Society." London, UK.

Pym, H. 2015. "Cancer, politics and public money - a difficult mix."

Radaelli, C. M. 1995. "The role of knowledge in the policy process." *Journal of European Public Policy* 2(2):159-83.

Raina, R. S. 1999. "Patronage and Evaluation in the Indian Council of Agricultural Research." *Evaluation* 5(3):278-302.

Rauschmayer, F., A. Berghöfer, I. Omann, and D. Zikos. 2009. "Examining processes or/and outcomes? Evaluation concepts in European governance of natural resources." *Environmental Policy and Governance* 19(3):159-73.

Read, C. 1995. *Preventing breast cancer: The politics of an epidemic*. London, UK: Pandora.

Rhodes, R. A. W. 1994. "The Hollowing Out of the State: The Changing Nature of the Public Service in Britain." *The Political Quarterly* 65(2):138-51.

—. 1997. *Understanding Governance: Policy Networks, Governance, Reflexivity and Accountability*. Buckingham, UK: Open University Press.

—. 2007. "Understanding Governance: Ten Years On." *Organization Studies* 28(8):1243-64.

Richards, D., and M. J. Smith. 2002. *Governance and Public Policy in the United Kingdom*. Oxford, UK: Oxford University Press.

Richards, M. 2010. "Review of the Cancer Reform Strategy." edited by Department of Health. London, UK.

Richardson, L. 2014. "Putting the Research Boot on the Policymakers' Foot: Can Participatory Approaches Change the Relationship between Policymakers and Evaluation?" in *Evidence and Evaluation in Social Policy*, edited by I. Greener and G. Bent. London, UK: Wiley Blackwell.

Risse, T. 2012. "Governance in Areas of Limited Statehood." in *The Oxford Handbook of Governance*, edited by D. Levi-Faur. Oxford, UK: Oxford University Press.

Rist, R. C. 1995. "Management Accountability: The Signals sent by Auditing and Evaluation." in *Policy Evaluation*, edited by R. C. Rist. Hants, England: Edward Elgar Publishing Company.

Rist, R. C., M. H. Boily, and F. Martin. 2011. *Influencing Change. Building Evaluation Capacity to Strengthen Governance*. Washington, D.C., US: The International Bank for Reconstruction and Development/The World Bank.

Ritter, T. 2007. "National Audit Office - the murky "clearance" process."

Ritz, A. 2011. "Attraction to Public Policy-Making: A Qualitative Inquiry into Improvements in PSM Measurement." *Public Administration* 89(3):1128-47.

Robert Francis QC. 2013. "Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry." London, UK.

Roche, C., and L. Kelly. 2012. "The Evaluation of Politics and the Politics of Evaluation." Developmental Leadership Program.

Rodríguez García, A. 2014. "Peña commands the creation of an area to overview quality of health services." in *Revista Proceso*. Mexico City.

Rogers-Dillon, R. H. 2004. *The Welfare Experiments: Politics and Policy Evaluation*. California, US: Stanford Law and Politics.

Romzek, B. S., and M. J. Dubnick. 1987. "Accountability in the Public Sector: Lessons from the Challenger Tragedy." *Public Administration Review* 47(3):227-38.

Rossi, P. H., and S. R. Wright. 1977. "Evaluation Research: An Assessment of Theory, Practice, and Politics." *Evaluation Review* 1(1):5-52.

Rotberg, R. I. 2004. "Strengthening governance: Ranking countries would help." *The Washington Quarterly* 28(1):71-81.

—. 2014. "Good Governance Means Performance and Results." *Governance* 27(3):511-18.

Rubí, M. 2014. "Coneval, sin consejeros y sin reglas para elegirlos ("Coneval, no counsellors and no rules for electing them")." in *El Economista*. Mexico City.

Rutkowski, D., and J. Sparks. 2014. "The new scalar politics of evaluation: An emerging governance role for evaluation." *Evaluation* 20(4):492-508.

Rutman, L. 1990. "Serving Management and Accountability: A Conflict for Program Evaluation? ." *Evaluation Practice* 11(1):39-43.

Rutter, J. 2012. "Evidence and Evaluation in Policy Making. A problem of supply or demand?", edited by Institute for Government.

Rutter, J., and M. Hallsworth. 2011. "Making Policy Better. Improving Whitehall's core business." London, UK: Institute for Government.

Sabatier, P. A. 1997. "Top-down and bottom-up approaches to implementation research." in *The Policy Process. A Reader*, edited by M. Hill. London, UK: Pearson Prentice Hall.

Sale, J. E M., L. H. Lohfeld, and K. Brazil. 2002. "Revisiting the Quantitative-Qualitative Debate: Implications for Mixed-Methods Research." *Quality and Quantity* 36(1):43-53.

Salisbury, C., K. Stewart, A. Cameron, S. Peckham, M. Calnan, R. Lart, S. Purdy, and H. Thorp. 2010. "Making the Most of Policy Evaluations: Overview and synthesis of evaluations of the White Paper 'Our health, our care, our say'." Bristol, UK: University of Bristol.

Sanderson, I. 2002. "Evaluation, Policy Learning and Evidence-Based Policy Making." *Public Administration* 80(1):1-22.

—. 2006. "Complexity, 'practical rationality' and evidence-based policy making." *Policy & Politics* 34(1):115-32.

Saunders, M. 2006. "Do We Hear the Voices?: The 'Presence' of Evaluation Theory and Practice in Social Development." *Evaluation* 12(2):251-64.

—. 2012. "The use and usability of evaluation outputs: A social practice approach." *Evaluation* 18(4):421-36.

Saunders, R. C., and C. A. Heflinger. 2004. "Integrating Data from Multiple Public Sources. Opportunities and Challenges for Evaluators." *Evaluation* 10(3):349-65.

Sawer, P., and L. Donnelly. 2014. "NHS scandal chiefs launch 'reputation management' firm." in *The Telegraph*. London, UK.

Sbragia, A. M. 2000. "Governance, the State, and the Market: What Is Going On?" *Governance* 13(2):243-50.

Schalock, R. L., and G. S. Bonham. 2003. "Measuring outcomes and managing for results." *Evaluation and Program Planning* 26(3):229-35.

Schedler, A. 2008. *¿Qué es la rendición de cuentas?* Mexico City.

Scheirer, M. A. 1998. "Commentary: Evaluation Planning is The Heart of the Matter." *American Journal of Evaluation* 19(3):385-91.

Schiavo-Campo, S. 2005. "Building Country Capacity for Monitoring and Evaluation in the Public Sector: Selected Lessons of International Experience." edited by The World Bank. Washington, D.C., US.

Schwandt, T. A. 2002. *Evaluation Practice Reconsidered*. New York, US: Peter Lang.

Schwandt, T., and P. Dahler-Larsen. 2006. "When Evaluation Meets the 'Rough Ground' in Communities." *Evaluation* 12(4):496-505.

Schwartz, R. 1998. "The Politics of Evaluation Reconsidered. A Comparative Study of Israeli Programs." *Evaluation* 4(3):294-309.

Schwarz, C., and G. Struhkamp. 2007. "Does Evaluation Build or Destroy Trust?: Insights from Case Studies on Evaluation in Higher Education Reform." *Evaluation* 13(3):323-39.

Scott, A. 2013. "Impact evaluation: which way forward? ." *The Guardian*.

Scriven, M. 1991. *Evaluation Thesaurus*. London, UK: SAGE.

—. 2007. "The logic of evaluation." Edgepress.

Seale, C. 2012. *Researching Society and Culture*. London, UK: SAGE.

Secretariat of the Interior. 2012. "Fifth National Survey about Political Culture and Citizen Practices." Mexico City.

Segerholm, C. 2003. "Researching Evaluation in National (State) Politics and Administration: A Critical Approach." *American Journal of Evaluation* 24(3):353-72.

Segone, M. (Ed.). 2008. *Bridging the gap. The role of monitoring and evaluation in evidence-based policy making*: UNICEF.

Selameab, T., and S. S. Yeh. 2008. "Evaluating Intangible Outcomes: Using Multiattribute Utility Analysis to Compare the Benefits and Costs of Social Programs." *American Journal of Evaluation* 29(3):301-16.

Shadish, W. R., T. D. Cook, and L. C. Leviton. 1991. *Foundations of Program Evaluation: Theories of Practice*. London, UK: SAGE.

Sharma, N. 2007. "Interactions and Interrogations: Negotiating and Performing Value for Money Reports." *Financial Accountability & Management* 23(3):289-311.

Shaw, I., J. C. Greene, and M. M. Mark. 2006. *Handbook of Evaluation: Policies, Programs and Practices*. London, UK: SAGE.

Simon, H. A. 1957. *Administrative Behavior: A study of decision-making processes in administrative organisation*. New York, US: Macmillan.

Simons, H. 2008. *Case Study Research in Practice*. London, UK: SAGE.

Søberg Shugart, M. 2006. "Comparative Executive-Legislative Relations." in *The Oxford Handbook of Political Institutions*, edited by R. A. W. Rhodes, S. A. Binder, and B. A. Rockman. Oxford, UK: Oxford University Press.

Solesbury, W. 2001. "Evidence Based Policy: Whence it Came and Where it's Going." ESRC Centre for Evidence Based Policy and Practice.

Sosa-Rubí, S., A. Salinas-Rodríguez, and O. Galárraga. 2011. "Impacto del Seguro Popular en el gasto catastrófico y de bolsillo en el México rural y urbano, 2005-2008 (Impact on the catastrophic expenditure of the *Seguro Popular* programme in the rural and urban Mexico, 2005-2008)." Mexico City: Instituto Nacional de Salud Pública.

Soss, J., and D. P. Moynihan. 2014. "Policy Feedback and the Politics of Administration." *Public Administration Review* 74(3):320-32.

Stake, R. E. 1978. "The Case Study Method in Social Inquiry." *Educational Researcher* 7(2):5-8.

Stame, N. 2006. "Governance, Democracy and Evaluation." *Evaluation* 12(1):7-16.

Steinich, M. 2000. "Monitoring and Evaluating Support to Decentralisation: Challenges and Dilemmas." European Centrre for Development Policy Management.

Stern, E. 2009. "Evaluation Policy in the European Union and its Institutions." *New Directions for Evaluation* 123:67-85.

Stern, J. 2010. "The Evaluation of Regulatory Agencies." London, UK: City University London.

Stevens, A. 2011. "Telling Policy Stories: An Ethnographic Study of the Use of Evidence in Policy-making in the UK." *Journal of Social Policy* 40(02):237-55.

Stufflebeam, D. 2001. "Evaluation Models." *New Directions for Evaluation* 2001(89):7-98.

Sturges, K. M. 2014. "External Evaluation as Contract Work: The Production of Evaluator Identity." *American Journal of Evaluation* 35(3):346-63.

Sullivan, H. 2011. "'Truth' junkies: using evaluation in UK public policy." *Policy & Politics* 39(4):499-512.

Supreme Audit Office. 2011. "Performance Audit. SICALIDAD.". Mexico, City.

—. 2012. "Health Quality System Performance Audit." Mexico City.

—. 2013a. "Financial Audit. Federal resources transfer through the agreement between the Ministry of Health and the State of Morelos." Mexico City.

—. 2013b. "Financial Audit. Federal resources transfer through the agreement between the Ministry of Health and the State of Zacatecas." Mexico City.

—. 2014a. "About us."

—. 2014b. "General Public Account Report 2013." Mexico City.

Supreme Auditor. 2015. "Delivery of the General Report of the Audit to the Federal Audit Account 2013." Mexico City.

Sutcliffe, S., and J. Court. 2005. "Evidence-Based Policymaking: What is it? How does it work? What relevance for developing countries?", edited by Overseas Development Institute.

Swinford, S., and C. Gammell. 2011. "Warnings of care home 'were ignored for months'." in *The Telegraph*. London, UK.

Syal, R., and M. Sweney. 2015. "Margaret Hodge accused of 'abusive and bullying' attack on BBC Trust chair." in *The Guardian*. London, UK.

Talbot, C. 2010a. "Performance in Government. The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom." Washington D.C., US: The World Bank.

—. 2010h. *Theories of Performance: Organizational and Service Improvement in the Public Domain*. Oxford, UK: Oxford University Press.

Talbot, C., and J. Wiggan. 2010. "The public value of the National Audit Office." *International Journal of Public Sector Management* 23(1):54-70.

Tansey, S. D. 2008. *Politics: The Basics*. London, UK: Routledge.

Target Ovarian Cancer. 2014. "General Election 2015 Manifesto." London, UK.

Taut, S., and D. Brauns. 2003. "Resistance to Evaluation. A Psychological Perspective." *Evaluation* 9(3):247-64.

Taylor, D. 2005. "Governing Through Evidence: Participation and Power in Policy Evaluation." *Journal of Social Policy* 34(04):601-18.

Taylor, D., and S. Balloch (Eds.). 2005. *The Politics of Policy Evaluation*. London, UK: The Policy Press.

Tépach, R. 2014. "Federal Budget Project for the Health Sector 2014-2015." Mexico City: Chamber of Deputies.

The British Monarchy. "What is constitutional monarchy?". London, UK.

The Conservative Party. "The Conservative Manifesto 2015."

The Health Foundation. 2014. "About the Francis Inquiry."

The John Hopkins University. 2006. "Monitoring Outputs and Outcomes And Introduction to Study Design."

The LSE GV314 Group. "Evidence-based politics. Government and the production of policy research."

—. 2014. "Evaluation under contract: Government pressure and the production of policy research." *Public Administration* 92(1):224-39.

The Patients Association. 2011. "Patients Association reacts to results of CQC." London, UK.

The Public Accounts Commission. 2008. "Corporate governance of the National Audit Office: response to John Tiner's Review." London, UK: House of Commons.

—. 2012. "Role." London, UK.

The Relatives & Residents Association. 2012. "Who we are."

—. 2013. "Care Quality Commission not yet successfully defined its core purpose or earned public confidence ". London, UK.

The Royal Institute of International Affairs. 2011. "Chatham House Rule." London, UK.

The World Bank. 1992. *Governance and development*. Washington, D.C., US: The World Bank.

—. 2004. *Monitoring & Evaluation: Some Tools, Methods & Approaches*. Washington, D.C.

—. 2008. "MX Results-based Management and Budgeting."

—. 2009. "Making Smart Policy: Using Impact Evaluation for Policy Making. Case Studies on Evaluations that Influenced Policy." Washington, D.C., US.

—. 2014. "A Model from Mexico to the World."

Themessl-Huber, M. T., and M. A. Grutsch. 2003. "The Shifting Locus of Control in Participatory Evaluations." *Evaluation* 9(1):92-111.

Thomas, G. 2011. *How to do your case study: A guide for students*. London, UK: SAGE.

Timmins, N. 2013. "Designed to fail: the Care Quality Commission is a warning to quango cullers." London, UK: Institute for Government.

Torrejón, C. S., C. Castillo-Durán, E. Hertrampf, and M. Ruz. 2004. "Zinc and Iron Nutrition in Chilean Children Fed Fortified Milk Provided by the Complementary National Food Program." *Nutrition* 20(2):177-80.

Transparencia Mexicana. 2011. "Índice Nacional de Corrupción y Buen Gobierno (National Corruption and Good Government Index) Executive Summary 2010.". Mexico City.

Trejo, Á. 2015. "Health Caravans arrive to Epazoyucan." in *Criterio*. Hidalgo, Mexico.

Trochim, W. M. K. 2009. "Evaluation policy and evaluation practice." *New Directions for Evaluation* 2009(123):13-32.

UK Aid, and Department for International Development. 2014. "Rapid Review of Embedding Evaluation in UK Department for International Development." London, UK.

UK Public Affairs Council. 2012. "Lobbying Definition."

United States General Accounting Office. 1990. "Case Study Evaluations."

Universidad Autónoma Metropolitana. 2012. "External Evaluation SICALIDAD." Mexico City: Ministry of Health.

Universidad Intercultural del Estado de México. 2013. "Health Caravans Evaluation Programme in High and Very Highly Marginalised Locations ". Mexico City.

Universidad Nacional Autónoma de México. 2012. "External Evaluation. Health Caravans Programme.". Mexico City.

University of Oxford. 2012. "Cancer costs the UK economy £15.8bn a year." Oxford, UK.

Vaessen, J., and F. L. Leeuw. 2010. *Mind the gap: Perspectives on Policy Evaluation and the Social Sciences*. New Jersey, US: Transaction Publishers.

Valovirta, V. 2002. "Evaluation Utilization as Argumentation." *Evaluation* 8(1):60-80.

Van Ackere, A. 1993. "The principal/agent paradigm: Its relevance to various functional fields." *European Journal of Operational Research* 70(1):83-103.

Van Der Heijden, J. 2014. "The Role of Government in Voluntary Environmental Programmes: A Fuzzy Set Qualitative Comparative Analysis." in *Public Administration*: Blackwell Publishing Ltd.

Van der Knaap, P. 1995. "Policy Evaluation and Learning: Feedback, Enlightenment or Argumentation?" *Evaluation* 1(2):189-216.

van Dooren, W., G. Bouckaert, and J. Halligan. 2010. *Performance Management in the Public Sector*. London,UK: Routledge.

Vanlandingham, G. R. 2011. "Escaping the Dusty Shelf: Legislative Evaluation Offices' Efforts to Promote Utilization." *American Journal of Evaluation* 32(1).

Vartiainen, P. 2002. "On the Principles of Comparative Evaluation." *Evaluation* 8(3):359-71.

Walshe, K., R. Addicott, A. Boyd, R. Robertson, and S. Ross. 2014. "Evaluating the Care Quality Commission's acute hospital regulatory model: Final Report."

Walshe, K., and D. Phipps. 2013. "Developing a strategic framework to guide the Care Quality Commission's programme of evaluation." London, UK: The University of Manchester.

Warrington Guardian. 2015. "CQC ratings to be displayed publically at health and care providers."

Watt, H. 2012. "Head of Care Quality Commission quits following damning report." in *The Telegraph*. London, UK.

Watt, N. 2011. "New GP cancer tests will save 5,000 lives a year, says Lansley." in *The Guardian*. London, UK.

Weaver, C. 2010. "The politics of performance evaluation: Independent evaluation at the International Monetary Fund." *Review of International Organizations* 5:365-85.

Weaver, K. R., and B. A. Rockman. 1993. *Do Institutions Matter? Government Capabilities in the United States and Abroad*. Washington, D.C., US: The Brookings Institution.

Weible, C. M., T. Heikkila, P. deLeon, and P. A. Sabatier. 2012. "Understanding and Influencing the Policy Process." *Policy Sciences* 45(1):1-21.

Weiss, C. H. 1973. "Where Politics and Evaluation Research Meet." *American Journal of Evaluation* 14(1):93-106.

—. 1993. "Politics and Evaluation: A Reprise with Mellower Overtones." *American Journal of Evaluation* 14(1):107-09.

—. 1998a. *Evaluation: Methods for Studying Programs and Policies*. UK.

—. 1998b. "Have We Learned Anything New About the Use of Evaluation?" *American Journal of Evaluation* 19(1):21-33.

Wells, P. 2007. "New Labour and evidence-based policymaking: 1997-2007." *People, Place & Policy Online* 1(1):22-29.

Wheeler, C. "EXCLUSIVE: Scandal of postcode lottery costing cancer patients' lives." in *Express*. London, UK.

White, F. 1999. *Audit, Accountability and Government*. Oxford, UK: Clarendon Press.

Widmer, T., and P. Neuenschwander. 2004. "Embedding Evaluation in the Swiss Federal Administration: Purpose, Institutional Design and Utilization." *Evaluation* 10(4):388-409.

Williams, D. D., and M. L. Hawkes. 2003. "Issues and Practices Related to Mainstreaming Evaluation: Where Do We Flow From Here?" *New Directions for Evaluation* 2003(99):63-83.

Wilson, C. A. 2013. *Public Policy: Continuity and Change*. USA: Waveland Press.

Winship, C. 2008. "Policy Analysis as Puzzle Solving." in *The Oxford Handbook of Public Policy*. New York: Oxford University Press.

Wood, M., and M. Flinders. 2014. "Rethinking depoliticisation: beyond the governmental." *Policy & Politics* 42(2):151-70.

World Health Organisation. 2006. "Quality of care: A process for making strategic choices in health systems.". France.

Wright, B. E., R. K. Christensen, and K. Roussin Isett. 2013. "Motivated to Adapt? The Role of Public Service Motivation as Employees Face Organizational Change." *Public Administration Review* 73(5):738-47.

Yin, R. K. 1992. "The Case Study Method as a Tool for Doing Evaluation." *Current Sociology* 40(1):121-37.

—. 2009. *Case Study Research: Design and Methods*. London, UK: SAGE.

York, N. 2013. "Developing an Evaluation Strategy-experience in DFID."

Zall Kusek, J., and R. C. Rist. 2004. "Ten Steps to a Results-Based Monitoring and Evaluation System." edited by The World Bank. Washington, D.C., US.

Zapico-Goñi, E. 2007. "Matching Public Management, Accountability and Evaluation in Uncertain Contexts: A Practical Suggestion." *Evaluation* 13(4):421-38.

Appendices

A. Main Differences between the Political Systems of the UK and Mexico

The institutional framework for evaluation and auditing in the UK and Mexico is influenced by the economic, cultural, social and political characteristics in which is embedded. Regarding the latter, it seems relevant to highlight the main differences that arise between these countries. The UK is a constitutional monarchy[1153] based on a parliamentary democracy, while Mexico is a democratic republic that operates within a presidential system. The differences between these constitutional designs have been pointed out by Søberg (Table A.1).

Table A.1 Main Differences between Parliamentary and Presidential Democracies

"Pure" Parliamentary Democracy	Presidential Democracy
1. Executive authority, consisting of a prime minister and cabinet, arises out of the legislative assembly; 2. The executive is at all times subject to potential dismissal via a vote of "no confidence" by a majority of the legislative assembly.	1. The Executive is headed by a popularly elected president who serves as the "chief executive;" 2. The terms of the chief executive and the legislative assembly are fixed, and not subject to mutual confidence; 3. The president names and directs the cabinet and has some constitutionally granted law-making authority.

Source: Based on Søberg Shugart, M., 2006. "Comparative Executive-Legislative Relations" in Rhodes et al., *The Oxford Handbook of Political Institutions*. Oxford University Press. P. 348

These differences matter from a comparative perspective for a number of reasons. One of the most transcendent for the purposes of this research is because it allows understanding how powers are allocated between the Legislative and the Executive. Also, because the political system also shapes how government rules and institutions are established. In this context, the work of Heady is relevant because he studied in detail political differences between the UK and developing countries, in which the Mexican case can fit. According to this author, in the UK (and also in the US):

> political culture and political structure are congruent, the political system is relatively stable and its legitimacy well established...

[1153] This refers to 'a form of government in which a king or queen acts as Head of State'. The British Monarchy. "What is constitutional monarchy?". London, UK.

> high ranking bureaucrats play substantial roles in governmental decision-making...
> External controls over bureaucracy are extensive and adequate in both Great Britain and the United States...
> Legislative oversight over administration relies in Great Britain on the doctrine of ministerial responsibility and makes slight use of specialized parliamentary committees...[1154]

In contrast, this author states that in developing countries it is perceived a contrasting reality ...

> Political instability is [a] prominent characteristic...
> Political leadership is concentrated in a minute segment of the population in most developing countries...
> The bureaucracies are deficient in skilled manpower necessary for developmental program [...] The shortage is in trained administrators with management capacity, developmental skills, and technical competence...
> Corruption, on a scale ranging from payments to petty officials for facilitating a minor transaction to bribes of impressive dimensions for equally impressive services, is a phenomenon so prevalent as to be expected almost as a matter of course...[1155]

In terms of bureaucracy, for instance, the differences exposed here show the existing culture and the way in which this element affects the adoption and internalisation of rules.[1156]

[1154] Heady, F. 2001. *Public Administration: A Comparative Perspective*. New York, US: Marcel Dekker. Pp. 198-210.

[1155] Ibid.Pp. 249-274.

[1156] Peters, B. G. 2001. *The Politics of Bureaucracy*. London, UK: Routledge.

B. Main Approaches to Policy Analysis

Analytical Approach	Basic Premises	Principal Authors	Advantages	Disadvantages	Is it suitable for this research?
Public choice	This approach derives from the rational choice theory and seeks to analyse public policy through the use of economic principles. The unity of analysis is the individual who is characterized as rational and self-interested beings. This approach studies the decision making process in the public sector focusing on variables such as uncertainty and asymmetric information. Public choice theory developed the concept of public goods as a result of the activity of public agencies.	Elinor Ostrom Vincent Ostrom Mancur Olson Douglass C. North	This approach allows understanding how public agencies work by studying the behaviour of individuals but mostly by identifying how individuals make decision in a context of uncertainty. In addition, public choice has provided a conceptualization of public goods and how the decisions are made within the public sphere and the impact on this kind of goods.	The economic approach limits the analysis of the whole dynamic of the policy making by ignoring other variables such as the effect of collective action.	This kind of approach is not suitable for this research because it doesn't provide enough theoretical tools to understand public action as a result of the interaction of different actors.
Garbage can model	This approach has as unit of analysis the organization. The basic premise is that within an organization, there is a set of problems and solutions, like in a garbage can, generated by the participants of the organization. The analysis made by this approach seeks to understand how a choice is made according to the conditions of the garbage can, the problems and solutions identified, the participants, the timing, among others.	Michael D. Cohen James G. March Johan P. Olsen	This approach provides a useful framework to understand how choices are made and how the context affects this decision. The garbage can model enables to understand how organizations work, particularly the way in which problems are categorised and solutions are allocated.	Even though the garbage can model attempts to analyse decision making within the organization, it fails in understanding and assuming the complexity of the collective action and the influence of external variables.	This model is not suitable for the research due to the fact that the research aims to understand in a more specific way how public problems are identified and how solutions are designed.

B. *Continued*

Analytical Approach	Basic Premises	Principal Authors	Advantages	Disadvantages	Is it suitable for this research?
Constructivist approach	The constructivist approach analyses how public policy is affected by the discourse, not only in a rhetorical way but also in other areas. This approach states that the discourse influence in all the process of policy making, particularly in the promotion of issues in the public agenda.	Jürgen Habermas Vivien Schmidt Joseph Schneider Giandomenico Majone	This approach is useful in explaining the role of discourse and the construction of arguments to achieve an objective in policy making. An approach of this nature is convenient due to the fact that there are political issues that affect policy making. Therefore, a constructivist approach helps to understand how the logic behind public policy is transmitted.	As a new approach the theoretical framework, as well as the methodology do not seem very clear. In addition, the analysis of the discourse seems limited to understand the policy making process as a whole.	The limits in the methodology of this approach, as well as its limits to understand relevant issues of the policy making process such as the institutional framework and the organizational context.
Policy networks Analysis	Policy network analysis focuses on the study of the interactions between the different actors that participate in the policy making. This approach seeks to explain how these actors are dependent on others in order to achieve their goals. In addition, this approach analyses the interaction within the context of the institutional framework and, consequently, how actors respond to the rules established.	R.A.W. Rhodes Tanja A. Börzel Maurice Wright	The policy networks analysis approach is a useful to understand a specific policy issue by observing and documenting the interrelation of the actors and their decisions during the process.	A disadvantage of this approach is basically that the complexity of the reality overpasses its premises, as the interrelations of the actors are not always open to the external observers.	This research focuses on the analysis of policy outputs and outcomes, even though the understanding of the interaction of the actors is important, this approach is not completely suitable for the research.

B.   *Continued*

Analytical Approach	Basic Premises	Principal Authors	Advantages	Disadvantages	Is it suitable for this research?
Organizational analysis Approach	The need for understanding how organizations work has become a recurrent issue in policy analysis. This approach seeks to study some organizational challenges and their implication in policy making. Some of these challenges are related to information problems, hierarchies, control, and communication. The study of public organizations has become a relevant topic within policy analysis, particularly after the implementation of strategies taken from the private sector.	Thomas Christensen Michel Crozier Niklas Luhmann Phillip Selznick	The study of organizations has become useful for understanding those elements that affect policy making. In addition, the organizational approach has developed different tools to analyse the behaviour of organizations, for example, the response to change and the adaptability.	The organizational approach is limited because it does not consider the external factors that affect policy making. Also, the study of organizations has focused on the implementation stage, leaving the rest of the stages less developed.	For this research, the organizational analysis is useful to understand some possible failures during the policy cycle, as well as some relevant issues regarding policy evaluation.
Policy cycle approach	This approach conceptualizes public policy as a series of "steps" or "stages", as a useful framework to understand how policy making works. In this cycle it is identified a problem to be solved followed by a series of actions directed to obtain a result to be assessed.	Wayne Parsons Peter Knoepfel Paul Sabatier	This kind of approach is useful for the identification of the actors and main issues of each of the stages, providing a deeper knowledge of how policy making works. In addition, this approach allows disaggregating the components of the process and enables the analyst to study in an individual way each of the elements of policy making.	The main disadvantage identified for this approach is that it is a rigid framework that might not suit the reality.	Due to the fact that the research focus on policy evaluation, the policy cycle approach is suitable for the research as it allows understanding the impact of every stage in the policy outputs.

C. List of Interviewees

ID	Description	ID	Description
UK-1	CQC Official	MX-1	Academician/Evaluator
UK-2	Health Committee Staff	MX-2	Ministry of Function Official
UK-3	NAO staff	MX-3	Ministry of Health Official
UK-4	NAO staff	MX-4	Ministry of Health Official
UK-5	Analyst NCIN	MX-5	Ministry of Health Official
UK-6	Researcher at Social Care Charity	MX-6	Ministry of Health Official
UK-7	Former Clerk PAC	MX-7	Ministry of Health Official
UK-8	DoH Official	MX-8	Ministry of Health Official
UK-9	DoH Official	MX-9	CONEVAL Official
UK-10	Former CQC Official	MX-10	Evaluator/Former CONEVAL Official
UK-11	Director at Cancer Charity	MX-11	Health Caravans Operator
UK-12	MP/PAC Member	MX-12	SAO Auditor
UK-13	NHS Official	MX-13	Researcher at Legislative Policy Analysis Centre
UK-14	Evaluator	MX-14	CONEVAL's Board Member
UK-15	Researcher at policy think-tank	MX-15	SICALIDAD Evaluator
UK-16	Evaluator	MX-16	Public Policy NGO Director
UK-17	Public Administration Select Committee staff	MX-17	SICALIDAD Operator
UK-18	Researcher at independent charity involved in government and public affairs	MX-18	Ministry of Finance Official
		MX-19	Ministry of Finance Official
		MX-20	Ministry of Finance Official
		MX-21	Researcher at Legislative Policy Analysis Centre
		MX-22	Former CONEVAL staff/Evaluator
		MX-23	Former Deputy
		MX-24	Evaluator/ Former CONEVAL Official

D. Performance Management Systems in the United Kingdom

Financial Management Initiative
Launched in 1982 with the white paper *Efficiency and Effectiveness;* underpinned by the concept of *Management by Objectives* first outlined by Peter Drucker in 1954 in his book 'The Practice of Management'. The initiative was designed to promote *inter alia:*
   (a) A clear view of the objectives, and the means to assess performance against them; and
   (b) Well defined responsibility for making the best use of their resources.

Next Steps Agencies UK
Launched in 1988 by Sir Robin Ibbs –Director of the Prime Minister's Efficiency Unit. The idea was to separate the policymaking work of government departments from the work of implementation, which was delegated to Departmental Agencies. Each agency set up with a Framework Document, agreed by the responsible Minister, setting out:
-Clear tasks
-Key Performance Indicators
-Annual Performance Reporting

Private Finance Initiative
This initiative […] involves payment rewards and penalties according to performance against targets throughout the period of the contract.

Public Service Agreement targets
Introduced in 1998, as part of the first Comprehensive Spending Review. Comprehensive Spending Review were designed to set out the resources each department would get over the next three years, and the targets that they would have to achieve with them. Public Service Agreement targets address the medium-to long-term priority outcomes of government policy. Each major government department agreed them with HM Treasury as part of each Spending Review.

Best Value Performance Indicators (BVPIs)
Introduced as part of the wider Best Value initiative under the Local Government Act 1999. Under the initiative, each local authority had to establish a set of authority-wide objectives and performance indicators. New performance targets were to be brought together with other service targets in the authority's annual Best Value Performance Plan.

Local Public Service Agreement targets
The Government announced in the Spending Review 2000 that it would introduce local Public Service Agreements (LPSAs). LPSAs were piloted with 20 authorities starting in late 2000 and a rollout to all upper tier authorities began in September 2001. Individual authorities sign up to targets to deliver key national and local priorities in return for operational flexibilities and financial rewards for success.

Comprehensive Performance Assessment

Announced in the December 2001 White Paper, 'Strong Local Leadership-Quality Public Services', the first Comprehensive Performance Assessments were conducted in December 2002. Comprehensive Performance Assessment is a standardised assessment framework that generates a single rating for each local authority. The two key elements of a council's activities that CPA seeks to measure are its core service performance and its ability to improve.

Performance Partnership Agreements
Performance Partnership agreements are Agreements between the Head of the Civil Service and the head civil servant of each department. They summarise the programme of civil service reform within each department –what is to be done, why and how; and how the centre will help.

Efficiency targets
Following the review by Sir Peter Gershon of public sector efficiency, the Spending Review of 2004 announced the introduction of departmental efficiency targets to realise the Government's ambition of achieving efficiency gains across the public sector on £21.5 billion and reducing 80,000 civil service jobs by 2008.

Local Area Agreements
Local Area Agreements represent a new approach to improve co-ordination between central government and local authorities and their partners, such as charities and non-governmental institutions, working through Local Strategic Partnerships. They focus on a range of agreed outcomes shared by all delivery partners locally.

Departmental Capabilities Reviews
Starting in October 2005, the Reviews are led by the Prime Minister's Delivery Unit and examine capacity within Department to achieve their objectives. The reports are to include clear assessments of current departmental capability and key actions to be taken to improve.

E. History of the National Audit Office

Early years	1314. Auditor of the Exchequer Earliest surviving mention of a public official charged with auditing government expenditure
	1559.Queen Elizabeth I established the Auditors of the Imprest
	1780. Commissioners for Auditing the Public Accounts were appointed by statute
	1834. Commissioners worked in tandem with the Comptroller, who was charged with controlling the issue of funds to the government
1860-1870	1860. First major steps towards proper financial accountability to Parliament were taken
	1859-1866. William Ewart Gladstone initiated major reforms of public finance and Parliamentary accountability
	1866. Exchequer and Audit Departments Act required all departments to produce annual accounts known as appropriation accounts, it also established the position of Comptroller and Auditor General (C&AG) and an Exchequer and Audit Department (E&AD)
	The 1866 Act established a cycle of accountability for public funds. The House of Commons authorizes expenditure. The Comptroller and Auditor General controls the issue of funds The results of the C&AG's investigations are considered by a dedicated parliamentary committee, the Committee of Public Accounts (PAC), established in 1861 by Gladstone
	1870. PAC took evidence from senior officials, normally Heads of Departments, who were designated as Accounting Officers by the Treasury
1921-present	1921. The 1921 legislation allowed the C&AG to rely in part on departmental systems of control and thus examine a sample of transactions, rather than all of them
	1960. Concerns were expressed by Parliamentarians and academics that the scope of public audit needed to be modernized to reflect the significant changes in the role of government over the course of the twentieth century
	1983. The National Audit Act 1983 was issued, under the Act, the C&AG formally became an Officer of the House of Commons, was given the express power to report to Parliament at his own discretion on the economy, efficiency and effectiveness with which government bodies have used public funds
	2000. Government Resources and Accounts Act 2000 introduced resource (accrual) based accounting and budgeting for Department Accounts
	2006. The Companies Act 2006 Established that the C&AG is able to compete for the audit of public bodies established as companies
	2011. The Budget Responsibility and National Audit Act 2011 Act established the NAO as a corporate body led by a Board consisting of four executive members (including the C&AG as Chief Executive) and five non-executive members (including a Chairman)

Source: National Audit Office, 2013. *History of the NAO.*

F. Key Findings of the Evaluation in Government Report

Coverage of Evaluation Evidence	Quality of Evaluation Evidence	Use of Evaluation Evidence	Production Resources and Barriers
Government guidance sets outs the expectation that all policies, programmes and projects should be subject to 'proportionate' evaluation; however, not all departments follow government requirements.	Departments' own assessment varies regarding the overall quality of their evaluation evidence.	Our review of the documents provided to HM Treasury by three departments during the 2010 Spending Review found limited references to evaluation evidence, which underpinned only a small proportion of resources that they sought from the Treasury.	Government departments use a wide range of models to commission and produce evaluations, but the rationale for this variation is not clear.
It is difficult to establish the coverage of evaluation evidence, but it does not appear to be comprehensive.	Our assessment of the fitness of purpose of a selection of 34 evaluations from four departments finds significant variation.	Impact assessments of policies under consideration rarely include relevant learning from evaluation evidence.	Information on staff time and budget spent on evaluation by departments is incomplete, so it is difficult for the government to take a view on whether the resources allocated are appropriate.
Departmental chief analysts recognise that gaps exist, but few departments have plans in place to evaluate all of their major projects.	We found some evidence that evaluation reports that are weaker in identifying causality tend to be more positive in assessing what the intervention achieved.	Public Accounts Committee and NAO reports have criticised departments for absent or poor-quality evaluation. There is little systematic information from the government on how it has used the evaluation evidence that it has commissioned or produced.	Independent evaluators outside the government experience difficulties accessing a range of official and administrative data that can be used to evaluate the impact of government interventions.

Source: National Audit Office, 2013. *Evaluation in government.*

G. The Cancer Strategy Reform (CRS) 2007

The Cancer Reform Strategy (CRS) developed presented by the Department of Health in 2007 derives from the NHS Cancer Plan issued in 2000. It is a five-year plan to improve services provided to cancer patients in the country. This document "sets out a programme of action across ten areas: six areas of action to improve cancer outcomes and four areas of action to ensure delivery" (Department of Health, 2007). These actions are listed below:

Actions to improve cancer outcomes
- Preventing cancer
- Diagnosing cancer earlier
- Ensuring better treatment
- Living with and beyond cancer
- Reducing cancer inequalities
- Delivering care in the appropriate setting

Drivers for delivery
- Using information to improve quality and choice
- Stronger commissioning
- Funding world class cancer care
- Building for the future by reassessing the progress made in tacking cancer

Source: Department of Health, 2007. *Cancer Reform Strategy.*

H. Key Findings of the Delivering the Cancer Reform Strategy Report

Improving the Quality of Information on Cancer	Strengthening the Commissioning of Cancer Services	Better Use of Resources
The National Cancer Intelligence Network (NCIN) has linked clinical, demographic and performance data from a range of sources and provided new analyses and reports on key cancer statistics, in addition to the development of a Cancer Commissioning Toolkit.	Many commissioners lack understanding of what drives costs and have not focused on improving value for money.	The Department has not monitored the cost of implementing the Strategy and the data it collects on expenditure on cancer services include unexplained variations from one year to the next within and between Primary Care Trusts (PCTs).
The Department has improved information on cancer by establishing a national baseline on levels of cancer awareness and undertaken a Cancer Patient Experience survey with responses from 67,000 patients.	Commissioners do not link cost and activity data to incidence, prevalence and survival data.	Significant reductions have been made in inpatient hospital bed days for cancer.
There is incomplete and inconsistent data on how advanced patients' cancers have become at the time they are diagnosed limit understanding of variations in outcomes and the effective allocation of resources.	Expenditure on cancer services is structured around complex payment mechanisms which largely fund activity in hospitals.	Poor coding of outpatient activity makes it difficult to measure follow-up activity after treatment.
There is duplication in the publication of cancer data which leads to confusion […] this leads to inefficiency in the publication of cancer data, and confuses commissioners about which data to use.	Chemotherapy and radiotherapy, two of the main treatments for cancer, are excluded from the national tariff and there is poor understanding of costs and activity.	The Strategy expected emergency admissions for cancer patients to be minimised and, while the rate of increase has been reduced, emergency admissions are still increasing.
	Commissioners have made progress in supporting the delivery of key commitments to improve cancer services, but shortcomings persist.	Challenging existing resource use can deliver savings, especially in the use of radiotherapy machines, in earlier diagnosis of patients, and in the reduction of length of stay for cancer patients.

Source: National Audit Office, 2010. *Delivering the Cancer Reform Strategy.*

I. External Scrutiny to the Care Quality Commission

Year	Scrutiny Body	Report	Main Findings
2011	Health Committee	*Accountability hearing with the Care Quality Commission*	1. The most striking and concerning aspect of the CQC's operation over the past year is the decision to divert resources from inspection and review activities towards meeting the demands placed on it by the process of registering providers, to the extent that the number of inspections fell far below usual levels. 2. The Committee concluded that the bias in the work of the CQC away from its core function of inspection and towards the essentially administrative tasks of registration represented a significant distortion of priorities. 3. We are extremely concerned that CQC's compliance activity fell to such low levels in the course of 2010-2011. 4. The long-standing vacancies for CQC inspectors are a further cause for concern. 5. The CQC should have identified the difficulties inherent in the regulations early in the registration process and made clear to the Government that unless modifications were made it would not be able to adequately to fulfil its duty to monitor and inspect providers. 6. The CQC must seek to address growing inspector caseloads through recruitment and should also bolster the support provided to inspectors to allow them to focus on their core frontline duties. 7. Quality and Risk Profiles have the potential to be a useful auxiliary tool for inspector, but in the present form the quality of data is limited in its reliability and coverage. 8. Action in the case of Winterbourne View was woefully inadequate: the CQC failed to 'actively follow-up' the local authority process, or conduct its own assessment, or even contact Mr Bryan for further information. 9. The CQC must accept responsibility for its poor handling of registration and adapt its processes accordingly. 10. The information currently provided by the CQC on adult social care providers is unhelpful and often out of date.

Source: Health Committee, 2012. *Annual accountability hearing with the Care Quality Commission.*

I. *Continued*

Year	Scrutiny Body	Report	Main Findings
2011	National Audit Office	*The Care Quality Commission: Regulating the quality and safety of health and social care*	1. The regulators for health and adult social care have been subject to considerable change in the last ten years. 2. The proposal to extend the Commission's role into new areas risks distracting the Commission from its core work of regulating health and adult social care. 3. There is a gap between what the public and providers expect of the Commission and what it can achieve as a regulator. 4. The Commission's budget is less than the combined budget of its predecessor bodies, although it has more responsibilities. 5. Responsibility for funding the regulation of health and adult social care is falling increasingly on the providers of these services rather than the Department. 6. The Commission underspent against its budget for 2009-10 and 2010-11, partly because it had a significant number of staff vacancies. 7. The timetable for registering health and adult social care providers, set by the Department, did not allow time for the registration process to be tested properly and the process has not run smoothly. 8. The Commission is seeking to learn lessons for the registration of GP practices, which has been deferred by a year. 9. Compliance review and inspection work fell significantly during 2009-10 and 2010-11. 10. The Commission has a systematic approach to assessing the risk that providers are not meeting the essential standards of quality and safety, but it depends on good quality information which is not always available. 11. The Commission has strengthened its whistleblowing arrangements in the light of the Winterbourne View case. 12. The Commission's performance management is constrained by gaps in data and reporting is mainly against quantity-based measures of activity.

Source: National Audit Office, 2011. *The Care Quality Commission: Regulating the quality and safety of health and social care.*

I. *Continued*

Year	Scrutiny Body	Report	Main Findings
2012	Public Accounts Committee	*The Care Quality Commission: Regulating the quality and safety of health and adult social care*	1. The Department is ultimately responsible for the effective regulation of health and adult social care but has not had a grip on what the Commission has been doing. 2. The Commission has been poorly governed and led. 3. The Commission's role is unclear and it does not measure the quality or impact of its own work. 4. The information provided to the public on the quality of care is inadequate and does not engender confidence in the care system. 5. The registration of GP practices must involve a meaningful assessment of compliance with the essential standards of quality and safety. 6. There are inconsistencies in the judgements of individual inspectors and in the Commission's approach to enforcement. 7. The Commission must strengthen its whistleblowing arrangements. 8. The Commission should not take on the functions of the Human Fertilisation and Embryology Authority at this time.

Source: Public Accounts Committee, 2012. *The Care Quality Commission: Regulating the quality and safety of health and adult social care.*

Year	Scrutiny Body	Report	Main Findings
2012	Health Committee	*2012 accountability hearing with the Care Quality Commission*	1. The CQC has been subject to sustained criticism in the period 2011-2012. 2. CQC Board member Kay Sheldon identified serious failings within the management, organisation, functions and culture of the CQC. 3. The CQC has developed a much keener focus on patient safety and has a better appreciation of what it exists to do, but we remain to be convinced that the CQC has successfully defined its core purpose. 4. The Committee concluded that the CQC's primary focus should be on ensuring that the essential standards it enforces can be interpreted by the public as a guarantee of acceptable standards in care. 5. There are indications that the CQC is developing its regulatory model to try and address this failing […] the inclusion of clinical expert advisors to support the inspection process is a positive step and we are confident that this will enhance the inspection process. 6. We were told that the CQC has learned from serious failings it had previously experienced in relation to whistleblowing […] we believe that the CQC can do even more to support the most vulnerable workers to come forward and report their concerns. 7. Primary care registration is a major challenge for the CQC and will test the degree to which the CQC has managed to implement learning from its previous experience with dental registration.

Source: Health Committee, 2012. *Annual accountability hearing with the Care Quality Commission.*

Year	Scrutiny Body	Report	Main Findings
2014	Health Committee	*2013 accountability hearing with the Care Quality Commission*	1. The CQC's new approach to registration and inspection was established in their revised strategy published in April 2013. 2. Lack of clarity and direction has previously undermined the CQC's attempts to establish itself as an authoritative regulator. 3. The Committee welcomes the fact that the CQC has now set out its objectives in clear terms. 4. The CQC has appointed three Chief Inspectors to oversee the inspection of Hospitals, Primary and Integrated Care, and Adult and Social Care. 5. The CQC has announced that it will now assess the quality of hospitals against a range of standards which are divided into three groups: Fundamentals of Care, Expected Standards and High Quality Care. 6. The evidence provided to the committee shows that the CQC has an ambitious timetable for reform registration and inspection […] however, some witnesses have expressed concern that the new model for registration and inspection will focus too closely on hospitals, primary care and social care at the expense of community, ambulance and mental health services. 7. The CQC plans that future inspections will be informed, and on occasion prompted, by an ongoing surveillance process. 8. The CQC plans to apply its methodology so it can provide an accurate assessment not only of the theoretical establishment of a provider but of the actual number of staff working. 9. The process of inspection is changing significantly as the CQC adopts a model of differentiated, in-depth inspection, the new model abandons generalist, generic inspections and from 2013-2014 the CQC intends only to operate "teams of inspectors who specialise in particular types of care". 10. In oral evidence David Behan provided an overview on how the CQC workforce will be restructured, telling the Committee that the role of every member of staff is likely to change, this will include creating specialist inspection teams and recruiting additional inspectors, a process that could take up to 18 months. 11. The CQC proposes a rating system for NHS hospitals comprising four separate classifications: Inadequate, Requires improvement, Good, and Outstanding. 12. The CQC's budget has risen rapidly in the last year and is likely to rise further […] it is no longer their objective to phase out grant-in-aid funding.

Source: Health Committee, 2013. *Annual accountability hearing with the Care Quality Commission.*