International journal of applied management and economics
Vol : 02 , N° 11 , December 2024
ISSN : 2509-0720

Assessing Public Policies: A Comprehensive Review of Evaluation Methods

- **AUTHOR 1 :** KOURA Abdelghani,
- **AUTHOR 2 :** BOUDHAR Abdeslam,
- **AUTHOR 3 :** OUDGOU Mohamed,

(1) : Laboratory of Research in Economics and Management of Organizations (LAREMO), National School of Business and Management, University Sultan Moulay Slimane, Beni-Mellal, Morocco..

(2) : Laboratory of Research in Economics and Management of Organizations (LAREMO), National School of Business and Management, University Sultan Moulay Slimane, Beni-Mellal, Morocco..

(3) : Laboratory of Research in Economics and Management of Organizations (LAREMO), National School of Business and Management, University Sultan Moulay Slimane, Beni-Mellal, Morocco..

Conflict of interest: The author reports no conflict of interest.

To quote this article: KOURA .A ,BOUDHAR .A & OUDGOU .M (2024) « Assessing Public Policies: A Comprehensive Review of Evaluation Methods »,

IJAME : Volume 02, N° 11 | Pp: 001 – 023.

Submission date : November 2024
Publication date : Decembre 2024

DOI : 10.5281/zenodo.14035228

International journal of applied management and economics
Vol : 02 , N° 11 , December 2024
ISSN : 2509-0720

Abstract

This paper aims to explore the theoretical foundations and methodological approaches of public policy evaluation, offering a detailed examination of how different evaluation types assess the effectiveness, efficiency, and impact of public interventions. By examining methods such as formative, summative, result-based, impact, ex-ante, and ex-post evaluations, we seek to understand the distinct roles each plays in shaping public policy outcomes across various domains. Through a comprehensive literature review, the article draws from multiple disciplines, including economics, political science, and public health, to provide a broader perspective on how evaluation frameworks are applied in different contexts. We highlight the challenges that evaluators frequently encounter, such as grappling with the complexity of public issues, the influence of historical decisions, and the political dynamics that often affect policy outcomes. Our literature review reveals that each type of evaluation serves a specific purpose in ensuring accountability, identifying unintended consequences, and measuring both the successes and shortcomings of public policies. This provides essential feedback to policymakers, allowing for continuous refinement and improvement of their strategies. Moreover, the article emphasizes the importance of rigorous protocols and objective assessments to guarantee credible and actionable results. Finally, we conclude by outlining critical areas for future research, such as improving the integration of evaluation methods, addressing evaluation biases, and enhancing the quality and transparency of ex-ante evaluations to further refine public policy development and implementation.

Keywords: Public policy, Policy Evaluation, Ex-ante Evaluation, Ex-post Evaluation, Impact Evaluation, Formative Evaluation, Effectiveness Evaluation, Process Evaluation

1. Introduction

At its core, policy analysis is like peeling back the layers of a complex system to see what drives decision-making (Koura, 2023; Kubler & de Maillard, 2009; Leca, 2012). It's about understanding the *"why"* behind the policies that shape our lives; why governments choose one path over another, who holds influence in these decisions, and how the realities of politics, economics, and institutions push and pull on the outcomes. Therefore, policy analysis lets us glimpse into the motivations and challenges policymakers face, helping us to see how decisions are crafted as responses to real-world problems (Howlett & Cashore, 2014; Strydom et al., 2010). But policies don't just stop at design; they come to life in the real world, where policy evaluation steps in. This is where the questions shift from *"why?"* to *"how well?"*. How well is a policy, or a policy mix working? Is it solving the problem it was meant to address, or is it falling short? Policy evaluation is like the moment of truth, it's when we measure the effectiveness of a decision by looking at its actual outcomes (Cashmore et al., 2010; Desplatz & Ferracci, 2016; Migaud, 2013; Penissat, 2011). Did it make a difference? Was it fair and efficient? Importantly, this isn't just a checklist of successes; it's about diving deeper, exploring unintended consequences, and learning from both the triumphs and the mistakes. It provides the critical feedback loop that helps improve or adjust policies over time (Perret, 2016).

Becker, (1985) claims that if the intentions of public policies were fully known, the public sector would prove to be a far more efficient producer and redistributor than generally believed. Indeed, many policies fail because they require the execution of tasks that are difficult to accomplish. Adding to this the propensity for corruption, incompetence, and political motivations to which these policies are subject, it seems quite normal that things often do not go as the policy-makers initially intended. Likewise, Lasswell, (1970) states that the specificity of policy evaluation lies in the fact that it is oriented toward problems that are not only caused by particular circumstances but also by historical decisions and constructions, whose composition can only be understood, through a return to and meticulous analysis of the past.

Given the plurality of governments' interventions, and consequently, public policies, the significant budgets allocated to them, and the sensitivity of the issues they address, evaluating their impacts is a major challenge for decision-makers and researchers, to measure their effects and identify the complications that hinder the achievement of their objectives. The evaluation of public policies as a scientific approach has been embraced by a plurality of disciplines such as education sciences, public health, sociology, political science, economics, and many other fields depending on the policies or programs being evaluated (Revillard, 2021). This phase of

the policy process falls within a broad spectrum of practices aimed at measuring and improving the performance of public institutions (Barbut et al., 2020). However, evaluation is always perceived as a delicate and difficult task, as many obstacles can distort and discredit an evaluation that does not adhere to a rigorous protocol (Langot & Petit, 2020).

The objective of this article is to explore the varied approaches to policy evaluation, tracing its evolution as a key element in the policy-making process. By reviewing the literature, this article will examine the different types of policy evaluation; from formative and summative evaluations to process and impact assessments, and how each plays a unique role in measuring the effectiveness, efficiency, and fairness of public policies. The rest of this paper is divided as follow; the second section sheds light on the controversial nature associated to the concept of policy evaluation and its taxonomy. The third section frames ex-ante and ex-post evaluations as part of a continuum from anticipating policy effects to assessing their real-world impact. Finally, the fourth section will present the most used approaches in policy evaluation, emphasizing the specificities, mechanisms, benefits and challenges related to each one.

2. Public Policy Evaluation: A Controversial Concept and Taxonomy

2.1 Defining the Core of Policy Evaluation

Traditionally, regarded as the concluding phase of the policy process, evaluation has often been overlooked in political science research. Nevertheless, it serves as a crucial tool for tracing the evolving dynamics within the political arena (Matyjasik, 2010). The evaluator's work lies somewhere between the role of a social sciences' researcher, who is interested in theory, research design, and data analysis (but who, for the most part, is not involved in service delivery), and the role of a practitioner, who deals with actors (people or institutions) in need, but is rarely interested or trained in data collection and analysis methods (Posavac, 2011).

This complexity of evaluation as a task, can also be seen when we try to define it; in fact, the concept of evaluation, presents challenges, as it encompasses diverse interpretations and theoretical frameworks that vary according to historical periods, contextual factors, and the perspectives of evaluators and scholars (Jabot, 2014). For instance, Luis and Moncayo, (2021) define evaluation, claiming its scientific status, as a science that systematically studies how, and with what success, interventions aimed at changing the world function. Another more synthetic definition is provided by *"La Société Française de l'Évaluation"* which defines evaluation as an activity aimed at producing knowledge about public actions, particularly concerning their effects, with the dual purpose of enabling citizens to assess their value and helping decision-makers improve their relevance, effectiveness, efficiency and convergence.

Arguably, policy evaluation has grown into a key part of the policy-making process. Once seen as an afterthought, it's now recognized as essential for understanding whether public policies work and how they can be improved (Davies, 2012). While the obvious weaknesses in the policy-making process represent enormous challenges, they are problems that can, in principle, be solved. More effort, more data, more transparency, better governance, competent experts, and goodwill. All of this can help to mitigate these problems and improve the implementation of public policies (Mueller, 2020). It is in this context that the evaluation of public policies emerges as a crucial tool to ensure the proper implementation, effectiveness, and efficiency of the policies and programs launched by governments.

Moreover, evaluation provides a foundation on which, decision-makers can make informed judgments about the effects that a public policy or a policy mix are likely to have on the group of beneficiaries to whom they are addressed, or on the conditions under which a desired effect is likely to be achieved or not. Indeed, studies show that policy-makers never start without benchmarks when they try to formulate new policies, but rather on the basis of a solid background from previous experiences.

It should be noted that, this polysemic and generalist notion is usually anticipated by another term that specifies its object (policy evaluation, program evaluation, process evaluation, practice evaluation, network evaluation, personnel evaluation…etc), its intended purposes (summative, formative, managerial, or democratic evaluation), the type of approach (collaborative, participatory, pluralistic evaluation), the type of inquiry (relevance, effectiveness, impact evaluation), the discipline or methods employed (socioeconomic evaluation, qualitative or quantitative evaluation), the mode of operation (external or internal evaluation), or the timing (ex-ante, intermediate, final, ex-post).

2.2 Taxonomy in turmoil

Over the years, contradictory perceptions of theorists and the varied methodological procedures to public policy evaluation have given rise, to a wide range of evaluative approaches, allowing evaluators to answer evaluation questions and refine their opinions based on the context, so that they can provide a perspective, that is close to and representative of the evaluated policy's effects.Palfrey et al., (2012) propose a productive tactic to categorizing public policy evaluations, by clarifying the primary objective pursued by each evaluation model and listing the different approaches adopted by practitioners and researchers, whose focus varies from one model to another (Table 1).

Table 1: Seven Models of Public Policy Evaluation

Model	Focus
Objective-oriented	The effectiveness, efficiency, and economic aspect of the evaluated intervention
Decision-oriented	Improves decision-making
Evaluation research	Provides explanations of results
Responsive	Evaluation process and participant perspectives
Goal-free	Openness to achievements other than those prescribed by the intervention's goals and objectives
Alternative-explanations	Alternatives to accepted descriptions of what is happening
Use-oriented	The usefulness of results for different stakeholders

Source: Palfrey et al. (2012)

Similarly, Crabbé & Leroy, (2012) propose a list of eleven approaches to evaluating environmental programs and emphasize that the choice of approaches is not random but depends on several factors, including the characteristics of the policy or program being evaluated, the evaluators and their perspective, and the evaluation's objective. Likewise, Stuart et al. (2017) advocate a reduced list of four approaches used in public policy evaluation, which we will further explain below. These approaches are (1) formative evaluation, (2) process or implementation evaluation, (3) outcome or effectiveness evaluation, and finally (4) impact evaluation.

Due to this variety of evaluation techniques, the evaluators must answer, one of the fundamental questions from the outset: "*What are we seeking to achieve through the evaluation?"* This question will allow them to set their expectations for the evaluation work. Are they looking to assess the quality of the service provided by the organization responsible for delivering the policy? Or are they seeking to enlighten potential improvements to the policy during its design or implementation? Or are they attempting to measure the results, or even the impacts, of a policy on its beneficiaries?

3. Ex-Ante and Ex-Post Evaluations: From Projections to Real-World Outcomes

3.1 Ex-ante Evaluation

The ex-ante method is conducted prior to the implementation of an intervention with the aim of providing, among other things, a preliminary estimate regarding the proper assessment of socioeconomic issues, the relevance of the strategy and objectives, and the coherence with other interventions or policies. Ex-ante evaluation seeks to quantify the expected effects of future

policies by considering the current state of a targeted area or population. This type of evaluation often involves simulations based on the behavior of the community or economy subject to the intervention.

In most cases, ex-ante evaluations rely on structural models containing the components of the environment faced by the stakeholders. In essence, they are *"what if?"* analyses. For instance, what would happen if certain characteristics of the tax system or public spending were altered? What would be the difference or impact on individual households compared to the initial situation, or status quo? Such analysis is marginal because it aims to capture differences relative to the status quo. Moreover, it is an almost necessarily behavioral, as it requires generating counterfactuals that account for the responses of agents (Bourguignon & Ferreira, 2003).

These structural models are based on underlying assumptions, which include identifying the main actors involved in the development of the policy in question, whether individuals, communities, public administrations, ministries, private actors, or even international stakeholders. Additionally, potential interactions between these actors and the different components of the environment are considered, as they will ultimately determine the potential outcomes (Koura et al., 2024). Furthermore, ex-ante evaluations are often conducted to provide decision-makers with information that enables them to allocate (accordingly) resources to activities that have a significant impact in terms of achieving the objectives pursued by their funders.

Helming et al., (2011) suggest that the process of conducting an ex-ante evaluation involves identifying the problem that triggers the need for the public policy. Subsequently, the objectives are set, and policy options that can be implemented are developed. Subsequently, the intended and unintended impacts of each option on the social, economic, and environmental variables of the system are defined, analyzed, and compared.

An insightful example of ex-ante evaluation is South Korea's study of its vehicle-to-grid (V2G) system, designed to assess potential benefits of integrating electric vehicles into the energy grid. This evaluation forecasted demand and examined how subsidies might influence electric vehicle adoption. By analyzing consumer interest and optimal pricing, it highlighted that a well-structured subsidy could promote social welfare, boost profits for V2G providers, and encourage electric vehicle use (Hong, 2012).

3.2 Ex-post Evaluations

Ex-post evaluations are conducted by decision-makers to assess the efforts made in the past to achieve previously set objectives, throughout all stages of the policy cycle and decision-making

context. These evaluations can be based on direct monitoring of key variables that reflect the objectives pursued by the policy, as well as comparing the achieved state or change against the expected impacts of the implemented policies. During the review phase of the policy cycle, the results of previously adopted policies can be compared to hypothetical or alternative counterfactual scenarios.

In fact, ex-post evaluations can offer valuable insights into how to adjust policy parameters to better achieve desired outcomes in the future or simply provide transparent reports on past policy performance. However, most ex-post evaluations struggle to derive insights that could not have been predicted through ex-ante evaluations. This challenge stems from the reliability and quality of the data available to the evaluator, as well as the burdensome task of collecting and processing such data, which can distort the findings of ex-post evaluations (OECD, 2004). To measure ex-post effects, an evaluation framework based on the OECD model is often used (Table 2). This framework aims to assess the various levels of success or failure of a public policy. It is a general framework that relies on both quantitative and qualitative data sources, encompassing the financial aspect of the evaluation through traditional cost-benefit analysis, and comparing outcomes with the budgets committed. Additionally, it includes supplementary measures of success or failure that can be important for decision-makers and other stakeholders (Meunier & Welde, 2017).

Table 2: Objective-Based Methodology for Ex-Post Evaluation

Success levels	Measures
Efficiency	Measuring operational success: To what extent are results achieved through the efficient use of financial, human and material resources?
Effectiveness	Measuring tactical success: Have the set objectives been achieved and to what extent has the project contributed to their attainment?
Impact	Measuring strategic success: Did the project have any positive or negative consequences other than those anticipated?
Pertinence	Measure of tactical success: Was the project consistent with the needs and priorities of decision-makers, intended users and other stakeholders?
Durability	Measuring strategic success: Are the positive effects of the project likely to continue after its completion?
Value for money	Has the project produced a positive net present value?

Source : Meunier & Welde, (2017)

Hypothetically, in a stable and predictable environment, ex-post evaluations should tend to align with ex-ante evaluations. However, due to various uncertainties and continuous fluctuations that economies experience, the observed outcomes from state interventions (ex-post evaluations) often diverge from the expected results (ex-ante evaluations). The degrees of convergence or divergence between the two evaluations primarily depend on the precision and quality of the ex-ante evaluation, as well as the fluctuations experienced by the environment and target population of the intervention, which significantly influence the output of the ex-post evaluation (Leite et al., 2011).

4. Key Approaches to Policy Evaluation

4.1 Formative Evaluation

When it comes to public policies[1], this category of approaches allows for checking the feasibility of a policy on several levels, whether it concerns the possibility of its implementation or the management of its potential social and economic effects. This type of evaluation is usually conducted during the design phase of the evaluated policy or when it is being modified in whole or in part. The term formative evaluation has been recognized since the 1960s, when researchers first made the fundamental distinction between the roles of formative and summative evaluation, identifying the goal of formative evaluation as the collection of information that can be used primarily for learning and intelligence purposes, impacting the ongoing development and improvement of the policy in question. In contrast, the primary objective of summative evaluation is to make a judgment and overall assessment of the program's effectiveness as a whole (Dehar et al., 1993).

In this category of evaluative approaches, the focus is more on the "operationalization" and implementation of the policy in question, and the quality of its development, organization, and implementation process. Formative evaluation, also known as developmental evaluation, aims to formulate, in advance, recommendations for making possible modifications and adjustments to the policy before its effective inauguration (Crabbé & Leroy, 2012). In this context, this approach reveals the discrepancies between what is feasible and what is potentially achievable, through a SWOT analysis, highlighting the strengths and weaknesses, opportunities, and threats associated with the implementation of the policy or program in question. Based on this analysis,

[1] In training engineering, formative evaluation refers to all the methods that a teacher or trainer uses to assess the learners' understanding, detect their learning needs, and assess their progress toward one or more previously set educational objectives. This allows the teacher or trainer (who takes on the role of an evaluator in this case) to adapt and modify their working methodology according to the learners' progress and reception.

the evaluator is expected to make suggestions for better implementation of the policy. Formative evaluation duly considers the unstable and dynamic political context and attempts to limit the policy's volatility, inherent in complex and varied projects in a politically unstable and dynamic environment.

Furthermore, Perret, (2016) asserts that the primary objective of formative evaluation is to improve the competencies and levels of stakeholder involvement, through an increased focus on process and quality, and by prioritizing the use of qualitative and participatory methods, enabling a clear view of the policy under evaluation. It is important to note that formative evaluation is always perceived as the opposite of summative evaluation. Indeed, these two evaluation methods are considered by Anglo-Saxons (particularly Americans) as archetypes. Although they do not have a univocal meaning in the field of evaluation, there is a consensus that formative evaluation aims to develop and optimize political programs, while summative evaluation is primarily concerned with collecting information on the effectiveness of policies. Formative evaluation is more action-oriented, while summative evaluation tends to be more research-oriented (Crabbé & Leroy, 2012). Another important characteristic of formative evaluation is that it involves a broader range of roles for the evaluator than those traditionally assigned to them. While the traditional role is limited to that of a technician or methodologist, formative evaluation involves an expansion of the primarily technical role, as well as political and consultative roles, such as those of educator, consultant, and change agent, all with the goal of improving the quality of the evaluation. Indeed, unlike traditional conceptions of the evaluator as a neutral and detached observer, formative evaluation requires the evaluator to work closely with the stakeholders involved in decisions related to planning, developing, and implementing the evaluated policy.

A formative evaluation of the E-Ready tool[2] showed how it could effectively support healthcare teams in launching new e-Health initiatives. By gathering feedback from healthcare providers through expert panels and interviews, the tool was refined to address real-world challenges, such as staff readiness and workplace adaptability. Through its structured, hands-on approach, E-Ready helped managers understand readiness levels, tailor support, and create action plans, ultimately making it easier for healthcare teams to embrace and successfully implement digital health solutions (Dannapfel et al., 2022).

[2] E-Ready is a self-assessment tool developed to help healthcare organizations gauge and improve their readiness for implementing eHealth initiatives by assessing factors like staff engagement, organizational support, and change readiness

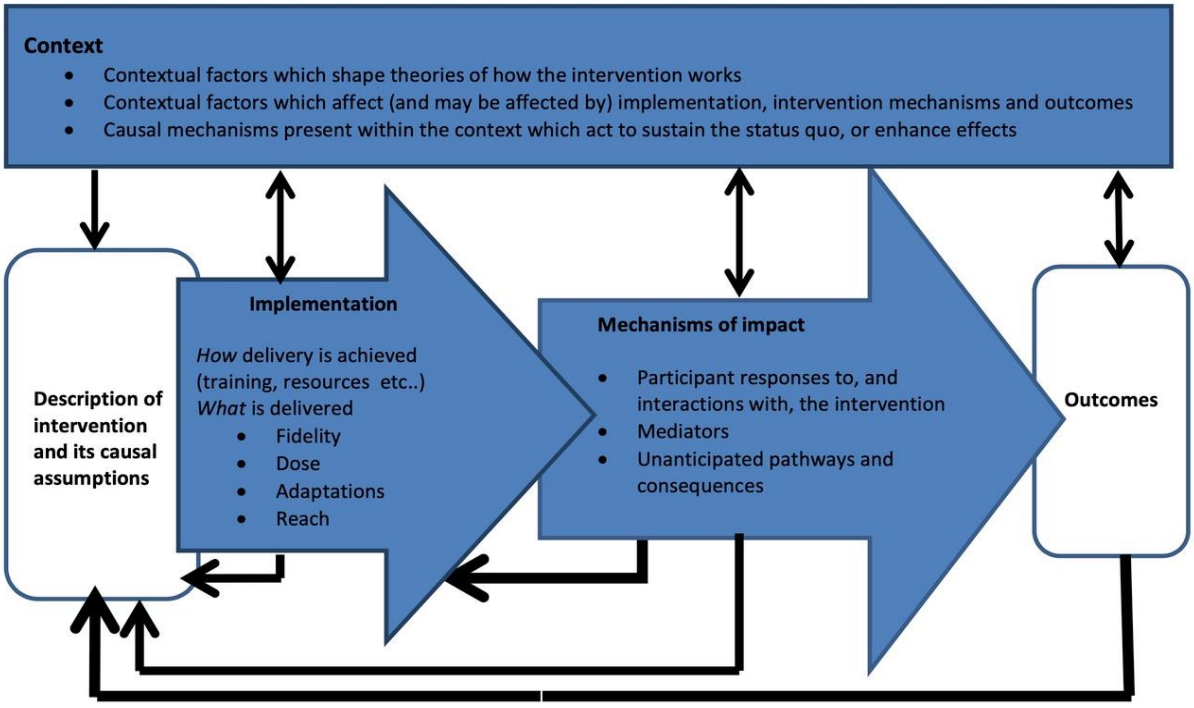
4.2 Process Evaluation

This type of evaluation analyses the managerial quality of a policy. It is often used with a decision-making logic to improve the implementation and delivery of the policy under evaluation. It does not primarily focus on determining whether the policy achieves its desired outcomes or impacts on its targeted population, but rather on assessing the quality of managerial decisions regarding the administration and procedures adopted for successful implementation of the policy. Dehar et al., (1993) conducted a study on process evaluations of health promotion and disease prevention programs, defining process evaluation as a method aimed at documenting and analyzing the operation of a policy to refine its outcomes and inform decision-makers for future planning of the policy.

Butterfoss, (2006) asserts that when it comes to public health policies, process evaluation is favored by evaluators and researchers to assess the scope, reliability, and quality of state interventions in health promotion and disease prevention. This method grants decision-makers the ability to refine concepts during implementation, allowing them to focus on the operational aspects of the policy and how its objectives will be achieved. Moore et al., (2005) argue that process evaluations have the primary goal of providing a detailed understanding, of the implementation of the policy under evaluation by analysing three main dimensions, as illustrated in Figure 1:

5. **Implementation**: Focusing on the structures, resources, and processes deemed necessary for implementation, as well as the quality and quantity of what was delivered;
6. **Mechanisms of impact**: This dimension examines how the policy and the interactions between the actors involved in its delivery affect its outcomes;
7. **Context**: This dimension evaluates how external factors influence the policy's outcomes and its implementation.

Figure 1: The Key Functions of Process Evaluation and the links Between Them



Source: Moore et al., (2005)

Overall, process evaluation emphasizes the type, scale, beneficiaries, and resources of the policy, as well as the practical problems encountered and how they are resolved. For effective process evaluation, planning must begin early to ensure the necessary data for evaluation is collected without compromise.

An illustration of process evaluation can be seen in the evaluation of Kenya's Free Maternity Services (FMS) policy. This evaluation focused on how effectively the policy was implemented, identifying gaps in funding mechanisms, reimbursement processes, and stakeholder engagement (Tama et al, 2017). Another example is seen in the urban river restoration project China, where treated wastewater was used to restore river flow. Through ongoing monitoring of nutrient levels and toxicity risks, the evaluation guided adjustments in restoration methods, such as filtration and aeration, to ensure water quality and prevent ecological issues like algal blooms (Sun et al, 2022).

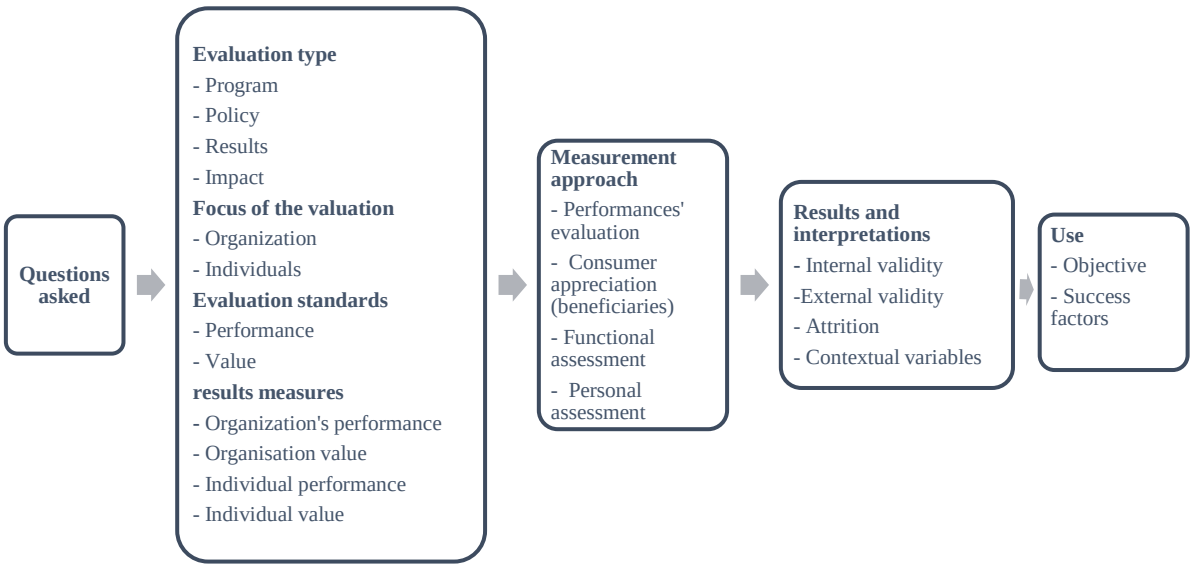
4.3 Results Evaluation

Kellaghan and Madaus, (2002) offer a comprehensive definition of results evaluation as he asserts that results evaluation is: *"A type of evaluation that uses person and organization-referenced outcomes to determine current and desired person and program-referenced outcomes and their use (program evaluation), the extent to which a program meets its goals and objectives (effectiveness evaluation), whether a program made a difference compared to*

either no program or an alternative program (impact evaluation), or the equity, efficiency or effectiveness of policy outcomes (policy evaluation)". This definition suggests that results-based evaluation primarily addresses the question of what the evaluated policy should accomplish for its beneficiaries, especially in terms of valued and referenced outcomes. It also considers the expected results that evaluators, stakeholders, and promoters, especially from the state, seek during the policy's design and implementation phases. These results are benchmarked by the organization to reflect its effectiveness and efficiency, using indicators and proxies to measure outcomes for the target population.

However, Cashmore et al., (2010) note that assessing effectiveness presents conceptual and methodological challenges, and despite growing interest and demand for evaluation, significant progress on these issues remains elusive. In fact, a persistent issue in evaluation research is the meaning of effectiveness itself, typically defined as the achievement of objectives but sometimes expanded to include cost-effectiveness. Nowadays, most public policies face two main evaluation needs: demonstrating increased accountability and ensuring continuous improvement. From a managerial perspective, this equates to results-based management, quality assurance, and outcomes valued by beneficiaries. Schalock, (2001) identifies five elements that should be included in results-based evaluation to meet these managerial needs involving two performance proxies and two value assessments: (1) performance evaluation, (2) consumer evaluation, (3) functional evaluation, and (4) personal appreciation (Figure 5).

Figure 1:The Elements Required to Implement a Results-Based Evaluation Approach



Source: Schalock, (2001)

Schalock (2001) adds that, in principle, results-based evaluation typically begins by asking questions. Although these questions can be multifaceted, five frequently asked questions by promoters, stakeholders, or evaluators directly relate to the four types of evaluation previously discussed in our work. These questions are:

- What outcomes does the policy produce for its beneficiaries?
- Does the policy achieve its goals (effectiveness evaluation)?
- Is the policy better than other options?
- Does this policy work?
- How can information about the outcomes be used to improve other policies?

The Pantawid Pamilyang Pilipino Program (4Ps) in the Philippines offers a strong example of results-based evaluation. This conditional cash transfer program regularly assesses its impact by linking financial aid to tangible health and education outcomes. Evaluations have consistently shown positive effects, such as increased school attendance and improved healthcare access for low-income families. This approach demonstrates how outcome-focused evaluations can guide program enhancements, ensuring that goals are met effectively and benefits reach those who need them most (Asian Development Bank, 2024).

4.4 Impact Evaluation

To illustrate the critical importance of impact evaluations, we will use an example from the health sector. Mali and Benin introduced a policy exempting payment for caesarean sections in 2005 and 2009, respectively. Ravit et al., (2018) attempted to evaluate the impact of this policy on the use of services and neonatal outcomes. The results show that the policy consisting of eliminating fees for caesarean section access, had a positive impact on delivery and significantly contributed to improving neonatal outcomes, particularly for the most disadvantaged women. These findings confirm the idea that user fees are a major barrier to caesarean section access in low-income countries. In addition to these findings, recommendations were made regarding the quality of government intervention, as the policy's implementation level was far from optimal, and its impact could have been greater if caesareans were truly free for all women.

Impact evaluations are part of a broader practice known as *"evidence-based policymaking"* (Leuz, 2018; Strydom et al., 2010). This practice, which has gradually spread internationally, is characterized by a shift in focus from inputs to outcomes and the results of each public intervention. This global trend is continually redefining how public policies are conducted. With a cost-effectiveness focus, results are increasingly used by decision-makers and demanded of them to enhance accountability, inform budget allocations, and guide policy decisions.

Gertler et al., (2016) provide a general and detailed definition of impact evaluation, he stated that: *« Impact evaluations are a particular type of evaluation that seeks to answer cause-and-effect questions. Unlike general evaluations, which can answer many types of questions, impact evaluations are structured around one particular type of question: What is the impact (or causal effect) of a program on an outcome of interest? This basic question incorporates an important causal dimension: we are interested only in the impact of the program, that is, the effect on outcomes that the program directly causes. An impact evaluation looks for the changes in outcome that are directly attributable to the program ».*

Indeed, an impact evaluation seeks to establish and quantify how an intervention affects the outcomes of interest to analysts and policymakers. Peersman, (2002) argues that, beyond the extent of the effects, an impact evaluation should contribute to identifying the actors and understanding how a program or policy succeeded, shedding more light on the determinants of the intervention's success or failure.

An effective impact evaluation should accurately assess the mechanisms by which beneficiaries respond to the intervention. These mechanisms may include connections to markets or improvements in social networks, as well as links to other existing policies. This latter connection is particularly important, as an impact evaluation that helps decision-makers understand the effects of an intervention can inform simultaneous and future impact evaluations of related interventions. The benefits of a well-designed impact evaluation are therefore long-term and can have considerable spillover effects (Shahidur et al., 2010).

For an impact evaluation to achieve its objectives, Peersman (2002) suggests that the framework and methodology used to analyze the meaning of the data collected must be systematic and transparent during the planning phase of the evaluation. The framework includes how data analysis will account for the assumptions made in the program's theory of change regarding how it is expected to produce the desired outcomes. In a truly mixed-methods evaluation, this includes using appropriate numerical and textual analysis methods, and triangulating multiple data sources and perspectives, to maximize the credibility of the evaluation's findings.

However, one of the major challenges facing impact evaluation is identifying a *counterfactual*, namely, the situation that would have prevailed without the policy or with an alternative policy. The fundamental problem of evaluation is that this counterfactual is, in practice, never observed. To address this, a variety of tools and instruments are used by evaluators and researchers. When it comes to policy impact evaluations, the various quantitative methods primarily aim to detect the impact (whether the objective is achieved or not), its source (the

notion of causality), and finally its magnitude, which is typically measured quantitatively. The primary focus is on assessing causality, particularly through the reasoning approach known as "counterfactual". The notion of causality was fully explained by the philosopher David Hume[3], who stated that causality cannot be directly observed but must be inferred from the observation of links between variables and the comparison of comparable situations (Baïz & Revillard, 2022).

Bourdin & Ragazzi, (2018) clearly define the principle of a counterfactual impact evaluation, emphasizing its reliance on a set of statistical techniques that allow evaluators to establish the extent to which the results obtained by individuals, companies, or regions that have benefited from a public policy are attributable to the policy itself, rather than to other factors that may affect these results. According to Bourdin & Ragazzi, statistics show a change in the objective variable without disentangling what is due to the effects of the policy from what is due to changes that would have occurred under the influence of other forces and factors (other than the policy). Unlike what happens in experimental sciences, we can only observe the change obtained by the units participating in the program; the counterfactual impact evaluation essentially aims to provide the best possible estimate of the (unobservable) evolution that these units would have achieved without the program and compare it to the actual evolution.

The concept of counterfactual is often illustrated through the question: *"What would have happened in the absence of the intervention?"* Consider two groups: the first, known as the treatment group, consists of individuals (citizens, communities, businesses, public institutions, etc.) who benefited from the public policy or policies under evaluation. The second group, known as the control group, includes individuals who did not benefit from the intervention. The counterfactual is derived by comparing the observed outcomes in both groups, which were initially identical and hypothetically would have remained so, if the intervention had not taken place within the treatment group. Hence, any difference between the two groups is attributed to the impact of the intervention.

To better understand this reasoning, let us express it through a simple mathematical model. The evaluator's goal is to measure the impact $\boldsymbol{\Delta}$ of an individual $\boldsymbol{i}$'s participation in a program on an outcome variable. Let $\mathbf{Y_1}$ represent the outcome an individual receives if they participate in the

[3] David Hume's approach to understanding causality is based on the idea that we can't directly see cause and effect. Instead, we learn about causality by noticing patterns—when one thing consistently follows another, we start to infer that there's a cause-and-effect relationship. This thinking is at the heart of how we evaluate policies today, especially when using the counterfactual method, which compares what actually happened with what might have happened if the policy hadn't been in place

program, and $\mathbf{Y_0}$ the outcome they receive if they do not ($\mathbf{Y_1}$ is unobservable for the individual if $\mathbf{Y_0}$ is observed, and vice versa). Consequently, for individual ***i***, the impact of participating in the program, Δ_i, is given by the following equation: $\Delta_i = \mathbf{Y_{1i}} - \mathbf{Y_{0i}}$. However, a significant challenge in counterfactual causal reasoning is the identification of treatment and control groups, which, as previously mentioned, must be identical in all aspects and remain so throughout the analysis period in the absence of the evaluated intervention. In practice, achieving such replication is impossible. Therefore, the evaluator seeks to assemble two groups that are as similar as possible, by identifying groups with convergent characteristics and behaviors.

Additionally, persistent differences observed during the analysis of the treatment and control groups allow for measuring the impact of the evaluated intervention. These measurements are often subject to uncertainty, typically expressed in quantitative methods through confidence intervals. For instance, the effect of a public policy promoting the financing of Moroccan SMEs on job creation could be expressed as follows: "[300,000; 500,000] with a 90% confidence level." This statement means that there is a 90% chance that the public policy contributed to the creation of between 300,000 and 500,000 jobs, and a 10% chance that the actual impact on job creation falls below 300,000 or above 500,000 jobs.

In qualitative methods, the reasoning around causality differs from that used in quantitative methods. Quantitative methods attempt to analyse the public intervention as a whole, primarily focusing on inferring the overall causality of the policy or policies in question. In contrast, qualitative methods aim to break down and analyse the intervention to describe the intermediate causalities observed at various stages of the policy process. Returning to the aforementioned example, the impact of a public policy promoting the financing of Moroccan SMEs on job creation can be approached using a quantitative method by inferring the causal link. For instance, this can be done by comparing employment in Morocco (considered the treatment group in this case) with employment in a similar country where the reform was not implemented.

The focus shifts from counterfactual causality to processual causality, also known as mechanical or physical causality, which involves critically analyzing the key causal links that connect the design of the policy to the achievement of its intended goals. This processual causality is based on a logic similar to that presented by the German philosopher (Carnap, 1973), who postulated: "*The cause of an event is not, strictly speaking, a thing, but a process. In everyday life, we say that certain things cause events. What we really mean is that specific*

processes or events cause other processes or events. Thus, we say that the sun causes the growth of plants; in fact, we mean that solar radiation, which is a process, causes this growth".

Following this logic, the evaluator focuses on providing insights regarding the quality of the policy's design and implementation, emphasizing aspects such as communication, coordination, monitoring, and convergence among the various parts of the ecosystem involved in addressing the issue targeted by the intervention being evaluated. Applying this processual causal reasoning to the same example, the evaluator might collect the perceptions of Moroccan SMEs managers regarding the program launched to provide them with the necessary financial resources to enhance their growth.

Moreover, the evaluator can also gather perceptions emanating from the different stakeholders intervening in the design, transmission and monitoring of the policy to assess the efficiency, coordination and especially, the convergence levels between the various actors of this ecosystem. This would allow the evaluator to observe, in the field, the reconciliations achieved in terms of both direct and indirect job creation. In turn, this would enable him to assess the public's perception of the policies and the quality of the communication established by policymakers (i.e., whether the target audience is informed or not), as well as the managerial aspects of the governmental action through the analysis of the perceptions of actors involved in the policy process.

Conclusion

Policy evaluation has become a cornerstone in modern policymaking, offering invaluable insights into how well public initiatives actually perform. Far from being an afterthought, evaluation is now a vital part of the policy cycle. It allows us to look beyond intentions and measure real-world outcomes, helping policymakers see where their strategies have succeeded, fallen short, or even led to unintended consequences. By using different types of evaluations—such as formative, summative, ex-ante, and ex-post assessments—governments and institutions can gain a clearer understanding of whether their policies are truly meeting the needs of the public.

One of the most important benefits of policy evaluation is its ability to create a feedback loop, giving decision-makers the data, they need to adapt and improve their policies. This process turns policy implementation into something dynamic and responsive rather than static. Instead of being locked into a single course of action, policymakers can adjust their strategies in real-time, based on solid evidence. Evaluation also plays a critical role in ensuring accountability. In an era where public resources are limited, it's more important than ever to make sure that policies are not just well-intentioned but also effective and efficient in their use of funds and effort.

However, as powerful as policy evaluation can be, it comes with its own set of challenges. The complexity of public issues, combined with the diverse approaches to evaluation, means that no single method fits all situations. Furthermore, political pressures and institutional dynamics can sometimes complicate the objective assessment of policies. Despite these hurdles, the need for rigorous and thoughtful evaluation has never been greater, especially as policies increasingly have to address interconnected, multifaceted problems.

Looking forward, there are several key areas where further research could significantly improve the field of policy evaluation. First, there is a growing need to better integrate mixed-methods approaches. While quantitative methods provide the hard data necessary for measuring outcomes, qualitative methods bring richness and context to the numbers, helping to explain why certain results occur. Future research should focus on developing models that blend these methods more seamlessly, offering a fuller picture of policy performance.

Another critical area for exploration is how to minimize bias in evaluations. Evaluators often face pressures from political stakeholders or institutional priorities, which can skew results. Developing frameworks that ensure neutrality and transparency in the evaluation process will help maintain the credibility and reliability of findings. As public policies increasingly cut

across multiple sectors, such as education, health, and economics, there is a pressing need to understand how to evaluate policies in these interconnected environments. Policies don't exist in isolation, and their impacts often ripple across different sectors. Future research should focus on creating evaluation models that can capture these broader, cross-sectoral effects.

Lastly, enhancing ex-ante evaluation methods is another area ripe for development. While ex-post evaluations help assess what happened after a policy was implemented, ex-ante evaluations could provide more robust insights into what might happen before a policy is enacted. By improving the tools and techniques used for ex-ante evaluations, policymakers can make more informed decisions from the outset, reducing the risk of unintended consequences later on.

Bibliography

Asian Development Bank. (2024). *Asian Development Outlook September 2024: Country Highlights*. Retrieved from https://www.adb.org

Baïz, A., & Revillard, A. (2022). Comment articuler les méthodes qualitatives et quantitatives pour évaluer l ' impact des politiques publiques ? Un guide à l ' usage des décideurs et des praticiens.

Barbut, L., Berriet-solliec, M., Péguy, T., Barbut, L., Berriet-solliec, M., Péguy, T., & De, L. (2020). L ' évaluation de politique publique : un exemple de pratique d ' amélioration des institutions publiques en tension entre efficience et pluralisme To cite this version : HAL Id : hal-02808782.

Becker, G. S. (1985). Public policies, pressure groups, and dead weight costs. Journal of Public Economics, 28(3), 329–347. https://doi.org/10.1016/0047-2727(85)90063-5

Bourdin, S., & Ragazzi, E. (2018). La science régionale et la performance des politiques publiques : retour sur les méthodes d'évaluation. Revue d'Économie Régionale & Urbaine, Février(2), 225–242. https://doi.org/10.3917/reru.182.0225

Bourguignon, F., & Ferreira, F. H. G. (2003). Ex ante evaluation of policy reforms using behavioral models. The impact of economic policies on poverty and income distribution, 123-142. 1–18.

Butterfoss, F. D. (2006). Process evaluation for community participation. Annual Review of Public Health, 27, 323–340. https://doi.org/10.1146/annurev.publhealth.27.021405.102207

Carnap, R. (1973). LES FONDEMENTS PHILOSOPHIQUES DE LA PHYSIQUE.

Cashmore, M., Richardson, T., Hilding-Ryedvik, T., & Emmelin, L. (2010a). Evaluating the effectiveness of impact assessment instruments: Theorising the nature and implications of their political constitution. Environmental Impact Assessment Review, 30(6), 371–379. https://doi.org/10.1016/j.eiar.2010.01.004

Cashmore, M., Richardson, T., Hilding-Ryedvik, T., & Emmelin, L. (2010b). Evaluating the effectiveness of impact assessment instruments: Theorising the nature and implications of their political constitution. Environmental Impact Assessment Review, 30(6), 371–379. https://doi.org/10.1016/j.eiar.2010.01.004

Crabbé, A., & Leroy, P. (2012). The Handbook of Environmental Policy Evaluation.

Dannapfel P, Thomas K, Chakhunashvili A, Melin J, Trolle Lagerros Y. A Self-help Tool to Facilitate Implementation of eHealth Initiatives in Health Care (E-Ready): Formative Evaluation. JMIR Form Res 2022;6(1):e17568 URL:https://formative.jmir.org/2022/1/e17568

DOI: 10.2196/17568

Davies P. (2012). The State of Evidence-Based Policy Evaluation and its Role in Policy Formation. National Institute Economic Review , 219, 41–52.

Dehar, M.-A., Casswell, S., & Duignan, P. (1993). Disease prevention programs. 17(2), 204–220.

Desplatz, R., & Ferracci, M. (2016). Comment évaluer l'impact des politiques publiques ? Un guide à l'usage des décideurs et praticiens. 72.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). Impact Evaluation in Practice, Second Edition. In Impact Evaluation in Practice, Second Edition. https://doi.org/10.1596/978-1-4648-0779-4

Helming, K., Diehl, K., Kuhlman, T., Jansson, T., Verburg, P. H., Bakker, M., Perez-Soba, M., Jones, L., Verkerk, P. J., Tabbush, P., Morris, J. B., Drillet, Z., Farrington, J., LeMouël, P., Zagame, P., Stuczynski, T., Siebielec, G., Sieber, S., & Wiggering, H. (2011). Ex ante impact assessment of policies affecting Land use, Part B: Application of the analytical Framework. Ecology and Society, 16(1). https://doi.org/10.5751/ES-03840-160129

Howlett, M., & Cashore, B. (2014). Conceptualizing Public Policy. Comparative Policy Studies, 17–33. https://doi.org/10.1057/9781137314154_2

JABOT, F. (2014). L'évaluation des politiques publiques Cadre conceptuel et étude de son utilisation par les décideurs des institutions régionales de santé en France.

Junhee Hong, Yoonmo Koo, Gicheol Jeong, Jongsu Lee, Ex-ante evaluation of profitability and government's subsidy policy on vehicle-to-grid system,Energy Policy,Volume 42,2012,Pages 95-104,ISSN 0301-4215,https://doi.org/10.1016/j.enpol.2011.11.053.

Kellaghan, T., & Madaus, G. F. (2002). 5. outcome evaluation.

Koura, A. (2023). La boite à outils de l'analyse des politiques publiques : une variété de modèles Public policy analysis toolbox: a variety of models Abdeslam BOUDHAR,) Mohamed OUDGOU. 4, 166–189. https://doi.org/10.5281/zenodo.10253062

Koura, A., Boudhar, A., & Oudgou, M. (2024). Convergence of public policies: Aligning policies and actors for sustainable financing of Moroccan SMEs—MACTOR analysis. Journal of Infrastructure, Policy and Development, 8(10), 6590. https://doi.org/10.24294/jipd.v8i10.6590

Kubler, D., & de Maillard, J. (2009). Analyser les politiques publiques. Analyser Les Politiques Publiques.

L. Schalock, R. (2001). Outcome-Based Second Edition.

Langot, F., & Petit, P. (2020). Évaluation Des Politiques Publiques. Revue Française d'économie, Vol. XXXV(1), 3–18. https://doi.org/10.3917/rfe.201.0003

Lasswell, H. D. (1970). The emerging conception of the policy sciences. Policy Sciences, 1(1), 3–14. https://doi.org/10.1007/BF00145189

Leca, J. (2012). L'état entre politics, policies et polity. Gouvernement et Action Publique, 1(1), 59. https://doi.org/10.3917/gap.121.0059

Leite, P., Narayan, A., & Skoufias, E. (2011). How Do Ex Ante Simulations Compare with Ex Post Evaluations?: Evidence from the Impact of Conditional Cash Transfer Programs. … of Conditional Cash Transfer …, 6(June), 1–43. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1871601

Leuz, C. (2018). Evidence-based policymaking: promise, challenges and opportunities for accounting and financial markets research**. Accounting and Business Research, 48(5), 582–608. https://doi.org/10.1080/00014788.2018.1470151

Luis, F., & Moncayo, G. (2021). Évaluation - Fondement, controverses, perspectives.

Matyjasik, N. (2010). L ' évaluation des politiques publiques dans une France décentralisée . Institutions , marché et professionnels To cite this version : HAL Id : tel-00543239.

Meunier, D., & Welde, M. (2017). Ex-post evaluations in Norway and France. Transportation Research Procedia, 26(2016), 144–155. https://doi.org/10.1016/j.trpro.2017.07.015

Migaud, D. (2013). Les cinq défis de l'évaluation. Revue Francaise d'Administration Publique, 148(4), 849–858. https://doi.org/10.3917/rfap.148.0849

Moore, G., Audrey, S., Barker, M., Bonell, C., Hardeman, W., Moore, L., Cathain, A. O., Tinati, T., Wight, D., & Baird, J. (2005). Process evaluation of complex interventions. UK Medical Research Council (MRC) Guidance.

Mueller, B. (2020). Why public policies fail: Policymaking under complexity. EconomiA, 21(2), 311–323. https://doi.org/10.1016/j.econ.2019.11.002

OECD. (2004). Tradeable Permits POLICY EVALUATION, DESIGN AND REFORM ENVIRONMENT.

Palfrey, C., Thomas, P., & Phillips, C. (2012). Evaluation for the real world: The impact of evidence in policy making. In Evaluation for the Real World: The Impact of Evidence in Policy Making (pp. 1–248). https://doi.org/10.3351/ppp.0008.0001.0008

Peersman, G. (2002). Overview: Data Collection and Analysis Methods in Impact Evaluation. Revista Latinoamericana de Psicologia, 34(3), 241–249.

Penissat, E. (2011). To quantify the "pure" effect of public action: between scientists battles

and redefinition of employment policies in France TT - Quantifier l'effet "pur" de l'action publique : entre luttes scientifiques et redéfinition des politiques d'emploi en France. Sociologie et Sociétés, 43(2), 223–247. https://search.proquest.com/docview/1171863347?accountid=14701 https://oculuo.primo.exlibrisgroup.com/openurl/01OCUL_UO/01OCUL_UO:UO_DEFAULT??url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&genre=article&sid=ProQ:ProQ%3Aibss&atitle=To+quanti

Perret, B. (2016a). L'évaluation des politiques publiques. Reflets et Perspectives de La Vie Economique, 18(1), 45–57. https://doi.org/10.3917/rce.018.0045

Perret, B. (2016b). L'évaluation des politiques publiques. Reflets et Perspectives de La Vie Economique, 18(1), 45–57. https://doi.org/10.3917/rce.018.0045

Posavac, E. J. (2011). Program Evaluation Methods and Case Studies x.

Ravit, M., Audibert, M., Ridde, V., De Loenzien, M., Schantz, C., & Dumont, A. (2018). Removing user fees to improve access to caesarean delivery: A quasi-experimental evaluation in western Africa. BMJ Global Health, 3(1), 1–12. https://doi.org/10.1136/bmjgh-2017-000558

Revillard, A. (2021). Quelle place pour les méthodes qualitatives dans l ' évaluation des politiques publiques ? To cite this version: HAL Id : hal-03458030.

Shahidur, R. K., Gayatri, B. K., & Hussain, A. S. (2010). Handbook on Impact Evaluation Quantitative Methods and Practices.

Strydom, W. F., Funke, N., Nienaber, S., Nortje, K., & Steyn, M. (2010). Evidence-based policymaking: A review. In South African Journal of Science (Vol. 106, Issues 5–6). National Research Foundation. https://doi.org/10.4102/sajs.v106i5/6.249

Sun, D., Lin, X., Lu, Z., Huang, J., Li, G., & Xu, J. (2022). Process evaluation of urban river replenished with reclaimed water from a wastewater treatment plant based on the risk of algal bloom and comprehensive acute toxicity. Water Reuse, 12(1), 1-10.

Tama, E., Molyneux, S., Waweru, E., Tsofa, B., Chuma, J., & Barasa, E. (2017). Examining the implementation of the free maternity services policy in Kenya: a mixed methods process evaluation. International Journal of Health Policy and Management, 7(7), 603–613.