



HAL
open science

Implementation Evaluation

Anne Revillard, Anu Rangarajan, Daniele Vidoni, Bruno Palier, Patrick Castel, Charlotte Halpern, Douglas J Besharov

► **To cite this version:**

Anne Revillard, Anu Rangarajan, Daniele Vidoni, Bruno Palier, Patrick Castel, et al.. Implementation Evaluation. Débats du LIEPP n°10, 2025, pp.13. hal-04952170

HAL Id: hal-04952170

<https://sciencespo.hal.science/hal-04952170v1>

Submitted on 17 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Implementation Evaluation

Introduction

Anne REVILLARD Sciences Po, CRIS, LIEPP

Before impact evaluation, process evaluation

Anu RANGARAJAN Mathematica

Driving Performance in the EU Context, the case of ERDF 2014-2020

Daniele VIDONI European Commission

Implementation as a dynamic process

Bruno PALIER Sciences Po, CEE, LIEPP

Organizations and power structures in implementation

Patrick CASTEL Sciences Po, CSO

Tackling complexity, including small-scale initiatives

Charlotte HALPERN Sciences Po, CEE, LIEPP

Concluding remarks

Doug BESHAROV University of Maryland

How to cite this publication:

Anne Revillard, Anu Rangarajan, Daniele Vidoni et al., **Implementation Evaluation**, *Débats du LIEPP n°10*, 2025-02-10.

Introduction

by Anne Revillard

As an interdisciplinary academic center for evaluation, one of LIEPP's goal is to make connections between analytical frameworks developed in the international field of evaluation and those developed in a diversity of academic disciplines. The question of implementation is typically an example of a process which has been the object of very separate theoretical and methodological developments in evaluation on the one hand, and in sociology and political science on the other.

On the one hand, reflections have developed in program evaluation around process or implementation assessment and implementation science, to go beyond a strict focus on impact measurement. These perspectives emerged as a way to complement impact evaluation by a better understanding of how impact is produced.

On the other hand, there is an abundant literature in sociology and political science on policy implementation. This body of literature was not developed in an evaluative perspective or driven by a preoccupation with the policy's impact. Yet this body of research, which has developed for decades, offers fruitful lessons about how the workings of implementation function, that can be most useful in an evaluative perspective. Yet the two bodies of literature have remained largely separated.

Drawing from the LIEPP seminar co-organized by LIEPP and the University of Maryland on June 5, 2024, this LIEPP Debates bring together contributions on policy implementation developed in evaluation, sociology and political science, based on the presentation and discussion of the Oxford Handbook of Program Design and Implementation Evaluation¹, coordinated by Anu Rangarajan.

The book is introduced by **Anu Rangarajan**, researcher at Mathematica, economist by training and coordinator of the publication. After giving an overview of the book's objectives, Anu Rangarajan develops three examples of implementation evaluation approaches: chronicle evaluation, compliance/fidelity evaluation, and translation evaluation. One of the book's contributors, **Daniele Vidoni** from the European Commission, then tackles the complexity of developing an evaluation framework for EU regional policy, combining outputs and result indicators.

Three leading Sciences Po experts in the study of implementation in sociology and political science then stress how the inputs from these disciplines can enrich implementation evaluation. After stressing how the EU has fostered a culture of evaluation, **Bruno Palier**, from the Centre d'Etudes Européennes et de politiques Comparées (CEE), points to process tracing as a useful tool of implementation evaluation. He insists on the role of contextual factors and the need to view implementation as « a dynamic process influenced by competing interests, cultural norms, and institutional capacities ». **Patrick Castel** from the Centre de Sociologie des Organisations (CSO), stresses the role of organizations and power dynamics in implementation. Finally, **Charlotte Halpern**, from CEE, draws on the case of environmental policies to show how evaluative practices may shape the policy landscape itself: by often focusing on the more visible actors, process evaluation may contribute to the marginalization of small-scale initiatives that have a transformative potential. She concludes by stressing the difficulty faced by political scientists studying implementation in being heard by policymakers who still tend to prefer straightforward impact measurement to more complexity-oriented process evaluation.

¹ Anu Rangarajan (ed.), *The Oxford Handbook of Program Design and Implementation Evaluation* (Oxford, New York: Oxford University Press, 2023).

Before impact evaluation, process evaluation

by Anu Rangarajan

In recent years, program evaluation within the United States and the international development sector has prioritized methods like randomized controlled trials (RCTs) to assess intervention impacts. This dominant focus on "what works" aims to provide robust evidence on the causal relationships between interventions and their outcomes. However, such evaluations often leave unanswered critical questions regarding the underlying reasons for success or failure. These include whether challenges arise from ineffective theories of change, contextual limitations, or implementation weaknesses. Addressing these gaps requires the adoption of a broader range of evaluation tools, particularly implementation evaluations.

Implementation evaluations offer a complementary lens that moves beyond outcome measurement to explore the processes and factors that drive or hinder program success. They are crucial for understanding not only why certain initiatives succeed or fail but also for determining how successful programs can be adapted and scaled in different contexts. Implementation evaluations encompass diverse methodologies aimed at addressing different stages of a program's lifecycle. They investigate questions such as: What were the key actions taken during program implementation? To what extent were these actions aligned with the intended design? How did contextual factors influence the outcomes? By answering these questions, implementation evaluations provide stakeholders with actionable insights for program improvement and scaling.

*The Oxford Handbook of Program Design and Implementation Evaluation*², authored by over 60

contributors from diverse professional backgrounds, provides not only theoretical insights but also practical recommendations derived from real-world experiences. In addition to chapters dedicated to various methodologies, the book includes case studies that illustrate these approaches in action. Its goal is to expand the toolkit available to researchers and evaluators, moving beyond the traditional focus on impact evaluations to embrace a richer diversity of techniques.

For this presentation, Anu Rangarajan draws on the chapter on process and implementation evaluation by Patricia Rogers and Michael Woolcock to highlight three primary types of implementation evaluations that are particularly in the context of impact evaluations: chronicle evaluations, compliance and fidelity evaluations, and translation evaluations.³ Each of these three approaches serves a distinct purpose in understanding program processes and outcomes.

Chronicle Evaluation: a detailed account of what happened during implementation

Chronicle evaluations aim to document the entirety of a program's implementation process, providing a comprehensive account of what occurred. They describe the program activities, the challenges encountered, and the ways in which the program's theory of change was translated into practice. Such evaluations are particularly valuable when examining novel or pilot programs, where little precedent exists to guide implementation.

A compelling example can be drawn from the education sector during the COVID-19 pandemic.

² Ibid.

³ Chapter 14 of the handbook by Rogers and Woolcock also discusses process approaches for answering causal questions (including process tracing, contribution analysis, and participant

attribution) as well as methods for answering evaluative questions about how good something is. Chapters 16-18 of the handbook discuss in detail implementation science evaluation approaches.

As schools closed for extended periods worldwide, governments and institutions adopted diverse strategies to ensure continued learning. In high-income countries, online learning platforms became a primary mode of instruction. However, in many low- and middle-income countries, where access to information and communication technologies (ICT) was limited, alternative approaches were necessary. These included broadcasting lessons via radio and television or distributing educational materials to parents for home-based learning. Chronicle evaluations in this context revealed stark disparities in access to educational resources, with poorer households often excluded from these solutions. Such findings highlight how emergency responses can inadvertently exacerbate inequalities, underscoring the need for careful program design that accounts for systemic inequities.

By documenting these dynamics, chronicle evaluations not only provide a clear understanding of what transpired but also generate lessons for future interventions.

Compliance/Fidelity Evaluation: whether intended activities were implemented appropriately

Compliance and fidelity evaluations focus on whether programs were executed as intended. They examine the alignment between planned activities and actual implementation, shedding light on variations in quality and consistency across different sites. These evaluations are particularly useful in identifying whether program failures stem from poor design or from weaknesses in implementation.

A large-scale agricultural project funded by the Millennium Challenge Corporation in Armenia offers an illustrative case. The program aimed to train 60,000 farmers, improve irrigation systems, and provide credit and post-harvest support. Initial impact evaluations revealed limited success, prompting a closer look at implementation processes. Process evaluations uncovered several issues: overly ambitious training targets led to the

inclusion of participants with little relevance to the program, such as family members of farmers attending sessions, additionally, several farmers reported attending training as they perceived this to be a means to get access to credit. Furthermore, critical program components like irrigation improvements and credit facilities were delayed, creating a mismatch between training timelines and the resources farmers needed to apply their new knowledge effectively. These findings not only explained the program's lack of impact but also provided actionable recommendations for improving the sequencing and targeting of future interventions.

Such evaluations demonstrate how implementation quality can significantly affect outcomes. By identifying gaps and inconsistencies, compliance process evaluations help refine programs and inform decisions about whether and how they should be continued or adapted.

Translation Evaluation: understand contextual factors needed to achieve positive outcomes in new settings or at larger scale

Translation evaluations examine how programs that succeed in one context can be adapted and scaled to other settings. These evaluations delve into the contextual factors—such as cultural norms, institutional capacities, and resource availability—that influence program effectiveness. By analyzing these elements, translation evaluations help determine the feasibility of replication and scaling efforts.

An example of this approach can be seen in the scale-up of a community-led sanitation program in Indonesia, originally piloted and demonstrated to be successful in Bangladesh. The program in Bangladesh, which used behavioral change approaches to improved sanitation, was driven by activities that evoked strong emotional responses, such as shame and disgust. However, when the initiative was implemented in East Java, Indonesia, researchers found that in many areas shame and disgust proved less effective due to differing cultural

perceptions of open defecation in the Indonesian context. Instead, health-related messaging, emphasizing the consequences of poor sanitation on child mortality and disease, emerged as a more effective motivator. Additionally, districts with strong governmental leadership and inter-agency coordination saw significantly better outcomes, underscoring the importance of local governance structures in program success. These insights enabled implementers to adapt the program for broader application, ensuring it resonated with the specific needs and values of new target populations.

Translation evaluations thus play a critical role in guiding scaling efforts, ensuring that interventions are not simply replicated but thoughtfully adapted to fit new contexts.

Conclusion

Implementation and process evaluations are indispensable tools for understanding and improving social programs. By focusing on the processes and contextual factors that underpin program outcomes, these evaluations provide invaluable insights for enhancing effectiveness and scalability. Whether through chronicling program activities, assessing compliance and fidelity, or facilitating the translation of successful interventions, these methodologies broaden the scope of traditional evaluation practices. They equip researchers, practitioners, and policymakers with the knowledge needed to design more effective programs and to maximize their impact across diverse contexts.

As evaluation practices evolve, the integration of implementation evaluations will be essential for addressing the complex challenges of social program delivery. These approaches not only deepen our understanding of "what works" but also illuminate the pathways to achieving sustainable and scalable success.

Driving Performance in the EU Context, the case of ERDF 2014-2020

By Daniele Vidoni

Evaluation has become a cornerstone of policymaking within the European Commission, serving as a tool to reflect on past initiatives and inform future actions. This is particularly true for regional policy, a flagship area of the European Union's cohesion strategy aimed at reducing disparities between regions. With a budget of €315 billion for the 2014–2020 programming period, regional policy represents nearly one-third of the EU budget, demonstrating the scale of its ambition. This contribution examines the evaluation mechanisms developed for this policy, focusing on the unique challenges posed by its size, multilevel governance, and shared management structure.

The EU's approach to regional policy evaluation reflects a dual necessity: balancing the overarching priorities set at the European level with the specific needs of 240 subnational regions. This balance required creating a flexible yet coherent system capable of capturing the effectiveness and efficiency of programs implemented across highly diverse contexts.

The Specificities of EU Regional Policy

Regional policy operates under a shared management system, where responsibilities are divided between the European Commission and member states. At the EU level, priorities and funding allocations are negotiated, while implementation is managed locally through operational programs. These programs vary significantly in size and complexity, ranging from projects worth a few million euros to those exceeding €5 billion, such as initiatives supporting small and medium-sized enterprises (SMEs) or cross-border cooperation.

The geographical diversity of regions also shapes the design of regional policy. Funding allocations are heavily weighted toward "less developed regions," defined as those with a GDP per capita below 75% of the EU average. This tiered approach ensures that resources are directed where they are most needed, though it also

introduces significant disparities in administrative capacities and access to data between regions.

Over time, the structure of regional programs has become increasingly articulated. Starting with simple project-based funding in the 1970s, programs have evolved to include specific priorities, predefined themes, and measurable objectives. This gradual refinement has allowed for a more strategic alignment between EU goals and local needs, but it has also created new challenges for monitoring and evaluation.

The 2014–2020 system of indicators for Regional Policy

The 2014–2020 programming period marked a turning point in the EU's approach to evaluation. The focus shifted toward developing a comprehensive monitoring system that could accommodate the complexity of regional policy. Two levels of indicators were adopted: output indicators, which measure the immediate products of investments (e.g., number of SMEs supported), and result indicators, which capture broader changes at the regional level (e.g., economic growth or job creation).

This indicator system reflects a deliberate choice to measure not only program outputs but also regional-level convergence. However, this approach presented significant challenges. Output indicators, while relatively straightforward to track, are limited in their ability to capture the broader impact of interventions. Result indicators, on the other hand, are conceptually valuable but difficult to attribute directly to specific actions, as they are influenced by numerous external factors.

A clear example of the EU's approach to evaluation can be seen in Tuscany's SME competitiveness program, where a portion of regional funding was allocated to the creation and consolidation of SMEs. This required Tuscany to collaborate with the EU to define specific activities and lines of support, as well as to establish a monitoring framework. Output

indicators, such as the number of new SMEs created, measured tangible results, while result indicators captured broader outcomes, including increases in SME size and total investments. This structured approach, applied across all EU programs, underscores the emphasis on linking actions to measurable results while accommodating the diverse contexts of regional policy.

Challenges for the 2021-2027 program

The start of the 2021-2027 program provided an opportunity to take stock of the 2014-2020 period and identify new challenges. On the positive side, the emphasis on evaluation fostered a culture of accountability and evidence-based policymaking. The requirement for managing authorities to design evaluation plans before program implementation encouraged forward-thinking and systematic data collection.

However, the sheer volume of evaluations conducted—over 1,800 during the programming period—highlighted disparities in evaluation capacity between regions. While some regions successfully leveraged evaluation findings to inform decision-making, others struggled to make use of the data, resulting in uneven impacts. Furthermore, the reliance on regional-level result indicators proved problematic, as it often disconnected evaluations from the practical realities of program implementation.

Looking forward, the EU has introduced a revised indicator system for the 2021–2027 programming period, incorporating three levels: outputs, direct results, and impacts. The inclusion of this new type of indicator reflects the desire to develop evaluation beyond the simple monitoring dimension. This new framework aims to address the limitations of the previous system by providing a clearer link between actions and their measurable effects.

Despite these advancements, the challenges of ensuring regional convergence remain significant.

The COVID-19 pandemic and the war in Ukraine have underscored the need for adaptable evaluation systems capable of responding to unforeseen crises. Maintaining a balance between accountability, flexibility, and inclusivity will be essential for the continued success of EU regional policy.

Implementation as a dynamic process

By Bruno Palier

Pragmatic Considerations on EU Evaluation Practices

The European Commission evaluation practices discussed reveal a gap when compared to established methodologies in social sciences and economics. Indeed, evaluations tend to focus on tracking the expenditure of funds and observable changes in the environment. However, the connection between financial input and observed outcomes remains tenuous. Rigorous gold-standard impact evaluations are rarely applied in these contexts due to their prohibitive costs and methodological demands, such as counterfactual analysis and long-term data collection.

The high costs associated with experimental methods, exemplified by projects requiring upwards of €300,000 for basic experimentation, deter policymakers from fully embracing impact evaluation. Nevertheless, the process of requiring governments and institutions to account for resource allocation and societal impact has fostered a cultural shift. For instance, EU funding mechanisms have significantly influenced state capacity and evaluation cultures, notably in countries such as Spain and Portugal, while also introducing evaluation practices in traditionally resistant contexts, such as France.

The pragmatic value of these evaluations lies not in providing definitive causal links but in promoting a culture of evaluation. Although the evaluations fall short of academic standards, they exert pressure on policymakers to substantiate the effectiveness of public spending, indirectly advancing public policy. However, this approach raises questions about the adequacy of current methodologies and the need for balancing financial constraints with methodological rigor.

Theoretical Reflections on Process and Impact Evaluation

A deeper theoretical discussion on evaluation methods highlights the divergence between process evaluation and impact evaluation. The former emphasizes the mechanisms and contextual factors linking inputs and outcomes, whereas the latter often simplifies these connections by assuming linear causality. Process tracing is a useful tool to identify these mechanisms. Rooted in qualitative single-case methodologies, this method advocates the importance of articulating and testing robust theories of change. By identifying the mechanisms that connect policy actions (X) to their outcomes (Y), process tracing provides a framework for understanding causality beyond simplistic associations.

Contextual factors also play a crucial role in evaluation, challenging the assumption of uniform applicability often inherent in impact evaluations. The drive to isolate variables and hold context constant overlooks the complex interplay of cultural, institutional, and historical factors that shape policy outcomes. Such omissions are particularly problematic in longitudinal studies, where ignoring temporal continuity and context diminishes the explanatory power of findings. This is particularly true for time-series analyses, which often overlook the influence of prior events on subsequent developments.

Finally, implementation studies reveal the inherently political and interpretative nature of policy execution. Implementation is rarely a straightforward translation of policy decisions but rather a dynamic process influenced by competing interests, cultural norms, and institutional capacities. Acknowledging this complexity, process-oriented methodologies advocate for integrating these dimensions into evaluation frameworks, thereby enhancing their relevance and explanatory depth.

Organizations and power structures in implementation

by Patrick Castel

Policy evaluation, particularly regarding health and social interventions, underscores a recurring dynamic where process evaluation often supersedes impact evaluation in practice. While policymakers and funders initially prioritize impact evaluations, practical constraints such as data inaccessibility, poor data quality, and the mismatch between evaluation timelines and political decision-making frequently redirect focus toward process evaluations. This shift highlights the critical role of understanding the organizational and relational dynamics inherent in policy implementation.

Historical and contemporary evaluations have consistently demonstrated that the success of implementation is intricately tied to the interplay of organizational objectives and power structures. Pressman and Wildavsky's foundational book *Implementation*⁴ illustrates mechanisms like conflicting objectives, actor-specific priorities, and temporal misalignments that impede the translation of policy decisions into effective practices. These challenges remain relevant, particularly in healthcare contexts where interventions, such as improving antibiotic use in nursing homes, encounter resistance from stakeholders whose priorities diverge from those of policymakers or, more broadly, "change entrepreneurs". For example, while public health officials and professionals emphasize antibiotic resistance, frontline professionals in nursing homes often prioritize immediate concerns like fall prevention and regulation of prescriptions other than antibiotics (anxiolytics, for example), revealing a fundamental misalignment in objectives (see **ORANEAT project**).

Further, Selznick's *TVA and the Grassroots*⁵ underscores the criticality of power relations in implementation. The dependency of implementing agencies on dominant resource-holding actors frequently results in deviations from original policy objectives. This relational conception of power—as an exchange of resources within an asymmetric but reciprocal dynamic—provides a valuable lens for evaluating public policy. In this framework, power stabilizes into structures that regulate behaviors, yet these same structures impose constraints on change and innovation, complicating the implementation and evaluation of policies⁶.

Organizations are both the facilitators and barriers of public policy implementation. While many policies aim to enhance coordination through the creation of new organizations, preexisting organizational ecologies often add layers of complexity⁷. This organizational density does not only complicate the implementation process but also affects the evaluation process itself. Evaluators and researchers, who must navigate intricate interdependencies to collect, share, and interpret data, often face significant cooperation challenges. Thus, understanding these power dynamics and organizational interfaces becomes imperative for meaningful process evaluations.

In some cases, process evaluation itself conflates with an evaluation of effectiveness, particularly in what has been termed "organizational policies." These policies focus on fostering inter-organizational cooperation to achieve collective goals, such as

4 Jeffrey L. Pressman et Aaron Wildavsky, *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 . . . Morals on a Foundation*, Third edition (Berkeley: University of California Press, 1984).

5 Philip Selznick et Jonathan Simon, *TVA and the Grass Roots: A Study of Politics and Organization* (New Orleans, La: Quid Pro, LLC, 2011).

6 Henri Bergeron et Patrick Castel, *L'organocène. Du changement dans les sociétés surorganisées* (Presses de Sciences Po, 2024), <https://shs.cairn.info/l-organocene--9782724643305>.

7 Patrick Castel et Léonie Hénaut, « Chapitre 9. Création organisationnelle et cercle vicieux: néo-bureaucratique », in *La société des organisations* (Presses de Sciences Po, 2022), 157-70, <https://doi.org/10.3917/scpo.borra.2022.01.0157>.

health, education, and security⁸. So, to evaluate how these new cooperative ventures are unfolding is to assess whether the policy has achieved its objective. A relevant example is the evaluation of precision medicine initiatives in oncology, where the aim was to accelerate translational research through the establishment of new organizational structures. These initiatives sought to bridge the gap between researchers and clinicians, fostering collaboration in areas like data sharing and joint clinical trials. Despite the intended synergy, significant challenges emerged, rooted in the unbalanced power dynamics among key stakeholders. For instance, pathologists, bioinformaticians, and computer technicians—actors often considered peripheral—played critical roles in implementing these projects. However, their perceived lower status within the hierarchy, compared to the influential key opinion leaders in the field, created friction and slowed collaborative efforts.

These findings underscore the necessity of integrating a sociological perspective into process evaluations. Beyond analyzing the formal organizational frameworks, evaluators must examine the relational dynamics and hidden power structures that can significantly influence policy outcomes. By doing so, they provide a more comprehensive understanding of the organizational complexities that shape the implementation and effectiveness of public policies. This approach necessitates significant investment in qualitative methods, such as extensive interviews and network analyses, to map the informal structures and power relationships that shape policy outcomes. While resource-intensive, these efforts are essential for addressing the complexities of modern public policy implementation.

8 Henri Bergeron et Patrick Hassenteufel, « Une contribution de la sociologie de l'action publique à l'évaluation de processus : Le cas des « politiques

d'organisation », *Idées économiques et sociales* 193, no 3 (21 septembre 2018) : 42-50, <https://doi.org/10.3917/idee.193.0042>.

Tackling complexity, including small-scale initiatives

by Charlotte Halpern

The evaluation of public policy implementation requires an acknowledgment of its inherently complex and fragmented nature. Implementation processes, as analysed in political science, remain unpredictable and prosaic, shaped by numerous decision points and contingent factors. These processes often involve competing goals, continuous bargaining among stakeholders, and outcomes that diverge from initial intentions. Implementation strategies frequently lack clarity, with public authorities and actors struggling to establish specific milestones or operational plans. Long-term goals, while ambitious, are often not supported by concrete steps to ensure their realization, leaving evaluators with the challenging task of reconstructing implementation pathways that are neither linear nor well-defined.

The foundational work of scholars like Lindblom or Wildavsky emphasizes that implementation is inherently messy, contingent upon numerous decision points, and shaped by the politics and conflicts of the moment. Policies often embody a symbolic nature rather than operational precision. A relevant example is environmental policies, which frequently served as declarations of intent rather than actionable frameworks, offering many opportunities for a variety of stakeholders to shape implementation processes. In the past, the absence of clearly defined goals and milestones made it difficult to measure outcomes or derive lessons from their implementation. As part of the climate transition agenda, new approaches have been developed to address these challenges. A wider variety of evaluation methods are being applied, such as process evaluation.

When addressing these challenges, it is crucial to consider how evaluative practices themselves shape the policy landscape. Process evaluation, while offering valuable insights, often privileges actors and

networks with established capacities, sidelining alternative approaches or smaller-scale initiatives that may hold transformative potential. For instance, in examining regional policy and international development, there is a tendency to rely on networks of actors already integrated into established programs. This focus excludes smaller, localized actors who might present innovative approaches to implementation but lack the visibility or institutional support to be included in evaluation programmes. The field of sustainable mobility planning within the European Union provides a useful illustrative case. Efforts to measure impact and foster partnerships often led to prioritizing pioneers and public-private collaborations while neglecting within public sector dynamics, governance capacities and smaller actors.

This brings us to the question of how evaluation results are received by sponsors, and their political implications. Indeed, the results of evaluations can call into question the very framework in which they are produced. In the case of sustainable mobility planning, this has meant confronting the limitations of centralized state mechanisms and questioning their efficiency in facilitating the processes required for transformative change.

These political questions are addressed to a greater or lesser extent depending on the methodological approach. Political scientists analyzing these processes often face challenges when presenting findings to EU and national authorities. Their perspectives, which highlight governance barriers and political dynamics, sometimes clash with the more outcome-driven approaches of colleagues in fields like economics, where evaluations often emphasize measurable impacts. While these economic approaches may yield robust quantitative results, they often overlook critical political and institutional factors that are essential to understanding and scaling successful policies.

Concluding remarks

by Doug Besharov

Why do social programs have such a poor track record? What can we do to maximize their chances of success? If a program shows promise, how can we scale it up? To answer these questions confronting policymakers, and to ensure that strong programs are designed, developed, implemented, and taken to scale, appropriate evaluation approaches must be used.

Anu Rangarajan is to be congratulated on producing a masterpiece of analysis and instruction. We often hear lip service paid to the combined importance of both implementation and impact evaluation. Her Oxford University Press Handbook on Program Design and Implementation Evaluation seems bound to reinvigorate the interactive process in research and practice. As she points out, implementation evaluation can determine "whether intended activities were implemented appropriately," "whether intended activities were implemented appropriately," and how to "understand contextual factors needed to achieve positive outcomes in new settings or at larger scale."

Dr. Rangarajan has recruited a team of eminent scholars to put together a comprehensive toolbox of evaluation methodologies that can be used to examine social programs throughout the stages of a program's life cycle. The methodologies presented in the Handbook describe how to conduct developmental evaluations, perform rapid-cycle evaluations, employ implementation science concepts, assess program effectiveness in the absence of a true counterfactual, measure cost-effectiveness, scale up promising interventions, and assess systems change. Incorporating these methodologies at every stage of a program's life cycle will maximize its chances of success; doing so consistently across social programs will help improve their track record overall.

She presented her main points at a meeting at Sciences Po's LIEPP followed by a series of commentaries, summarized above. I was most taken by the resonance of the comments across disciplines and continents. Their comments attest to the immensity of the challenges before us, but, as well, that the outcome will surely be worth the effort. Thanks to LIEPP for presenting such a formidable program.



Le LIEPP (Laboratoire interdisciplinaire d'évaluation des politiques publiques) bénéficie du soutien du plan d'investissement France 2030 à travers l'IdEx Université Paris Cité (ANR-18-IDEX-0001).

www.sciencespo.fr/liepp

Si vous voulez recevoir les prochains échos du LIEPP et rester informés de nos activités, merci d'envoyer un courriel à : liepp@sciencespo.fr

Directrice de publication :

Anne Revillard

Edition et maquette :

Andreana Khristova

Evane Grossemy

Sciences Po - LIEPP
27 rue Saint Guillaume
75007 Paris - France
+33(0)1.45.49.83.61



Distributed under a Creative Commons [Patrimoine - Partage selon les Conditions Initiales](https://creativecommons.org/licenses/by-nc-sa/4.0/) 4.0 International License