



Speech

Strengthening evaluation capacity in Government: Why institutionalisation matters and why it is not sufficient*

Evaluation

2023, Vol. 29(2) 161–167

© The Author(s) 2023

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/13563890231166981

journals.sagepub.com/home/evi



Daniele Lamarque

European Evaluation Society; European Court of Auditors, France

If we are interested in promoting and strengthening evaluation in a country, we know that the extent to which evaluation is institutionalised is key. Investigating the strengths and weaknesses of institutional factors and identifying their limits help us understand under which conditions these factors will make a difference – and what is still needed. I will be drawing on several studies and reports – as well as my own experience – that address the issue of how evaluation is institutionalised internationally. While all these sources confirm the positive influence of institutional arrangements, they also show that evaluation develops unevenly and slowly and these developments are often difficult to specify or measure.

Assessing how far evaluation is institutionalised

The International Atlas of Evaluation (Furubo et al., 2002) provided the first systematic comparative overview of evaluation culture and practice across 21 OECD (Organisation for Economic Co-operation and Development) countries. This overview was updated 10 years later for 19 countries (Jacob et al., 2015). This update concentrated on nine indicators also used in the first Atlas to measure the maturity of evaluation in a country and compare different national evaluation cultures. These indicators give a good overview of what it means when we say evaluation is institutionalised:

*This is an edited version of a Keynote address delivered to a Workshop on the 'Institutionalisation of Public Policy in Portugal' organised in Lisbon by PlanAPP and OECD in November 2022.

Corresponding author:

Daniele Lamarque, European Court of Auditors, Paris, France.

Email: daniele.lamarque@orange.fr

- Evaluation takes place in many policy domains – there are regular evaluation activities across various policy fields;
- There is a supply of evaluators specialising in different disciplines – evaluators from different academic disciplines have mastered different evaluation methods and provide advice over evaluations;
- Discussions and debates fuel a national evaluation discourse;
- A national evaluation society exists – evaluators have their own societies, networks or frequently attend meetings of international societies that discuss evaluation standards or ethics;
- Institutional arrangements exist in government for conducting evaluations and disseminating their results to decision-makers;
- There are institutional arrangements in Parliament for conducting and disseminating evaluations to decision-makers;
- There is evaluation pluralism within each policy domain: different people or agencies commission and perform evaluations;
- Evaluation activities occur within the Supreme Audit Institution (SAI) – and these SAIs either conduct evaluation activities themselves or oversee conditions for undertaking evaluations within the public sector: maybe even carrying out different forms of meta-evaluation;
- Evaluations focus on outcomes as well as inputs/outputs; public sector evaluations need to demonstrate that public interventions actually have impacts on the problems they were intended to solve.

The results reported in the updated Atlas were broadly positive. It showed that in the intervening years, the majority of countries had improved their practices and achieved a higher degree of maturity of evaluation culture.

More recently, other studies and reports have also addressed evaluation capacity from an institutional perspective. These studies and reports have shown that there continues to be discernible, if uneven, progress in many countries. For example, Stockmann et al. (2020) include reviews of the institutionalisation of evaluation in 16 countries plus the European Union (EU) and also provide an extensive bibliography on the topic. Systematic information and comparisons can also be found in various OECD reports (see, for example, OECD, 2020).

Evaluation in the executive branch: A prominent but compromised role

Most studies highlight the prominent *role of the executive branch* – what OECD calls the Centre of Government (COG) – in promoting the institutionalisation of evaluation. This indeed is the focus of the 2022 Recommendation of the OECD Council on public policy evaluation (OECD, 2022).

Having the primary initiative of designing and implementing policies, the government holds several instruments for facilitating the development of evaluation. It can lay down regulations or guidelines on evaluation, insert evaluation clauses in new regulations, issue *ex ante* assessments, develop performance measurement in budgetary documents, ensure coordination, develop training and introduce quality assessment systems.

Although *ex ante* evaluation and regulatory impact assessment (RIA) are not defined as evaluation as such, they can play a significant role in enhancing the overall evaluation cycle. The EU Commission experience shows how they can contribute to designing the intervention logic, identifying policy options and their relative impacts, and preparing for future monitoring and evaluation. RIAs can also determine what indicators should be used, who will be responsible for gathering which data, against what benchmarks of success or failure, and when an evaluation should take place. The key components of RIAs for preparing future evaluations are: problem definition and use of evaluation, objectives and intervention logic, impacts, future monitoring and evaluation.

However, although two-thirds of countries surveyed by OECD have adopted some kind of legal framework and three-quarters have adopted evaluation guidelines, the institutionalisation of evaluations as part of the responsibility of the government remains difficult to achieve. A wide range of explanations of this limited progress have been put forward:

- If not internalised or incorporated into the culture of administrations, legal obligations may be neglected or implemented in a formal way that does not influence decision-making.
- Embedding evaluation in the budgetary process is key for the development of sound management. It is necessary to increase efficiency and accountability and usually requires the involvement of Parliament and civil servants in the process. However, the OECD, among others, has noted limited progress in such ‘embedding’. Explanations offered include different timescales – evaluations are rarely conducted on an annual basis – and irrelevant performance indicators – not sufficiently related to outcomes.
- Varone and Jacob (2004) analyse different contextual and institutional factors that may have an impact on the institutionalisation of evaluation, and consequently on its culture. They found that ‘New Public Management’ reforms that prioritise efficiency and rationalisation may weaken evaluation. For example, addressing budgetary deficits leads to using evaluation as a tool for justifying cuts.
- Embedding evaluation in the decision-making process and fostering evaluation feedback have to be connected to a strong evidence-informed policy system and centres of expertise – which are not always present.
- Sectoral evaluation systems facilitate professionalisation (e.g. in development aid, education or other domains) but may lead to evaluator-capture conducting ritual evaluations or sometimes self-appraisals.

The ‘whole-of-government approach’ recommended by OECD allows a more comprehensive framework for evaluation, able to encompass a series of interlinked policies and to address the systemic challenge of complexity. It should be noted, however, that this framework, and most studies on institutionalisation, does not take into account subnational levels (or do so only marginally). This gap is particularly damaging considering that the great majority of public policies are implemented at different levels of government, with a strong involvement of local entities.

Institutionalising evaluation depends on a complex partnership

Although governments play a significant role, evaluation cannot be only dependent on their sole action. Developing evaluation capacities involves a complex partnership of different institutions or stakeholders acting collectively or at least interacting with each other.

Parliaments, as key partners of the executive for policy-making and part of the democratic process, are often cited as potential strong actors in evaluation. However, it must be noted that parliaments play a rather weak role in evaluation in Europe (with the exception of Switzerland).

This situation is far from being satisfactory because of the parliamentary dual function in the legislative process and for the control of the executive. Parliaments should help to improve the law-making process, foster public performance, and hold the government to account. They can work with the support of the SAI or academia.

The main reason of this weakness probably lies in the lack of time and expertise, as well as the absence of real political will.

SAIs are encouraged to perform evaluations (INTOSAI GOV 9400, 2016), and a number of SAI have extended their performance audit activities into the evaluation sphere. They can be active players in this domain, provided that they clarify their intentions, distinguishing, for example, between evaluation and compliance or performance audit. They also need to adopt relevant organisational and methodological arrangements. SAI can contribute both by performing evaluations on their own initiative or at the request of Parliament and by checking how evaluation works within the government (as suggested by OECD).

There is no evidence that strengthening institutionalisation of evaluation necessarily leads to improvements in result utilisation and quality. Regulations may help, but formal obligations to evaluate cannot replace both individual and collective commitment to adopt an ethic of effectiveness in public administration. This of course requires that administrations abandon what is often a legalistic rather than an effectiveness culture. We have learned that all institutional actors implicated in the framing and delivery of evaluation have to work in cooperation with each other, if they are to ensure a sound and sustainable development of evaluation in the public sector.

Institutionalisation affects all the objectives that evaluation sets out to achieve

I have suggested that the institutional framework of evaluation is made up of several actors, public and private, vested with different responsibilities and functions throughout all policy processes that underpin public action. The institutional framework therefore has to be analysed as an ecosystem where interrelations are at least as important as individual roles, and institutional culture is as crucial as institutional responsibility. I have also argued that any institutional framework needs to be supported by a political will and a common awareness of performance issues. All actors have their limits and weaknesses, but they are stronger when they cooperate.

This issue can be more deeply understood in relation to the key objectives of evaluation, objectives that are continuously presented as pivotal although scarcely achieved.

- The *contribution of evaluation to improving the decision-making process* cannot be achieved without a strong connection between relevant actors be they political or administrative, fulfilling responsibilities of policy design, adoption and implementation. This raises the sensitive question of the role of expertise and knowledge to inform both decision-makers and evaluators, and facilitate evidence-informed policy-making. The role of academia, the contribution of different disciplines, the balance between quantitative and qualitative approaches, and their relevance to the problems to assess

are all relevant here. However, it is also essential to bear in mind that policy objectives are per se ambiguous and multiple, and that power resists most attempts to be rationalised or disciplined (Gilbert and Verrier, 2016).

- The *quality of evaluations*, a condition for their credibility and acceptance, and their capacity to build a strong culture and practice, appears to be a weak point in most countries (OECD, 2022). This shortcoming refers to the capacity of experts to enter the field of meta-evaluation; trigger debates between academia, practitioners and other partners; and promote ethical and professional evaluation standards. The role of networks, public debates, training and academic courses, events held by national evaluation societies, should not be minimised. International and national networks are crucial.
- The role of evaluation in *strengthening democracy* is universally espoused but still difficult to define and fulfil. Experience and studies confirm the weak contribution of civil society promoting the emergence and the use of evaluations, despite the fact that participatory approaches are a prominent part of evaluation activities and research.
- Evaluation is expected to *reinforce cultural change* in government and address the growing *complexity of public policies*. The presentation of the institutional framework gives the impression of a rational and well-organised landscape with clear roles devoted to each actor. The contribution of evaluation to encouraging cultural change is limited partly because the reality of evaluation's institutional framework is not as rational or well-organised as we would like it to be.

Part of the explanation of why evaluation does not consistently meet its objectives stems from the discrepancies between the espoused discourse and observed practice among most institutional actors in the evaluation space. It would certainly be difficult to find any actor in any country that would not claim that public action should be result-oriented, open to public consultation and regularly measured. However, practices often diverge from this consensual affirmation. Many administrations remain focused on regulations, and auditors prefer compliance audit to performance audit, since the former is easier to perform, more familiar and more attractive for the media. Politicians recommend evaluation, but are more reluctant when it comes to evaluating their own decisions, and remain convinced that the most appropriate evaluation is the election.

The discourse about evaluation as such (its methods, its merits, its standards) remains mainly concentrated within specialised networks, evaluation societies, academia or symposia. Many civil servants are able to speak of performance and adapt their management accordingly while still ignoring evaluation. Evaluation issues are rarely visible to the general public, through either traditional media or social networks. And when such issues are discussed, evaluation is perceived as a technocratic activity, or as a scientific approach that can be contradicted (which is normal), or, which is worse, as non-reliable given the complexity of real-world policy-making and implementation.

The importance of multi-level governance including the local level

Among the most urgent of these real-world complexities is the *evaluation of shared or multi-level policies*. These remain a grey area. Most assessments of the institutional framework of evaluation applies primarily to central government and does not take into account subnational entities present in federal states or decentralised countries. This gap should be bridged if

evaluators want to respond to emerging issues, such as the specificities of territories and the different and often contradicting expectations of citizens towards public policies – like environment, mobility or agriculture, for example. Similarly, the evaluation of European programmes is too often undertaken by specialised units linked to the Commission, while such evaluations can be quite disconnected from other national, regional or local policy evaluations.

It appears increasingly necessary that evaluations address local issues, notwithstanding the legal and institutional constraints. Three main categories of evaluations serving local needs can be identified:

- Local actors need to evaluate their development projects and strategies;
- They need to know about the effects that interventions planned and implemented by central actors produce on their territory;
- They need to address the challenge of coordination, as we know from studies of multi-level governance, that many local problems involve actors with overlapping responsibilities that intersect at the local level.

Conclusion

Despite definite but uneven progress, limited institutionalisation continues to hold back evaluation culture and practice. Changes in the way the executive arm of government organises evaluation are not enough. Many more partners have to engage. Evaluators have, more than ever, to encourage the development of evaluation within governments, Parliaments and SAIs. Evaluation societies play a key role and deserve more resources and active support from their members and from academic communities. The public debate about how evaluations are conducted and utilised must go on.

However, new challenges are arising, involving actors, strategies, territories or sectors of society, which require specific approaches – all of this at a time when a revolution in the data on which evaluators depend is well underway. Key questions raised by the studies on institutionalisation should inspire national models taking into account historical, political and cultural characteristics, at different levels of our multi-level governance systems. These need to be tailored to national specificities although working with common institutional frameworks so as to maintain comparability between national systems.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

References

- Furubo JE, Rist RC and Sandahl R (eds) (2002) *International Atlas of Evaluation*. New Brunswick, NJ; London: Transaction Publishers.
- Gibert P and Verrier PÉ (2016) Peut-on discipliner le pouvoir? Étude sur le contrôle de la rationalité managériale par le juge dans trois innovations législatives françaises. *Politiques et Management Public* 33(3–4): 165–96.

- INTOSAI GOV 9400 (2016) Guidelines on the evaluation of public policies. Available at: <https://www.eca.europa.eu/sites/evaluation-public-policies-programmes/en/Documents/INTOSAI%20GOV%209400%20-%20Guidelines%20on%20the%20Evaluation%20of%20public%20policies%20-%20Benjamin%20Negre.pdf>
- Jacob S, Speer S and Furubo JE (2015) The institutionalization of evaluation matters: Updating the International Atlas of Evaluation 10 years later. *Evaluation* 21(1): 6–31.
- OECD (2020) *OECD Public Governance Reviews: Improving Governance With Policy Evaluation – Lessons From Country Experiences*. Paris: OECD Publishing.
- OECD (2022) Recommendation of the Council on public policy evaluation (OECD/LEGAL/0478). Available at: <https://strata.gov.lt/wp-content/uploads/2022/10/Recommendation-on-Public-Policy-Evaluation.pdf>
- Stockmann R, Meyer W and Taube L (eds) (2020) *The Institutionalisation of Evaluation in Europe*. Palgrave Macmillan, Switzerland.
- Varone F and Jacob S (2004) Institutionnalisation de l'évaluation et nouvelle gestion publique: Un état des lieux comparatif. *Revue Internationale De Politique Comparée* 11(2): 271–92.

Daniele Lamarque is Past President of the European Evaluation Society and was between 2014 and 2019 a member of Cour des comptes européenne (European Court of Auditors). She was a member of the French Court of Auditors (1984–2013) and chair of two Regional Audit Chambers (1997–2005 and 2011–2013) and participated in several institutional systems set up after 1990 for the development of public policy evaluation in France.