

---

# From Studies to Systems: Ray Rist's Influence on Evaluation Systems: Insights from International Research Group for Policy and Program Evaluation (INTEVAL)

Maria Bustelo  
*Complutense University of Madrid*

Steve Jacob  
*Université Laval*

*Journal of MultiDisciplinary Evaluation*  
Volume 21, Issue 50, 2025

**JMDE**  
Journal of MultiDisciplinary Evaluation

ISSN 1556-8180  
<http://www.jmde.com>

**Background:** Evaluation systems aim to embed evaluation as a standard and routine practice within public organizations. By integrating evaluation into the everyday operations of government, these systems have the potential to enhance the relevance and use of evaluation findings, support organizational learning, and contribute to more transparent and accountable governance. Although evaluation systems are often promoted as tools for strengthening the connection between evaluative knowledge and decision-making, it is important to examine how these systems may both facilitate and constrain the development and meaningful use of evaluation.

**Purpose:** This article seeks to draw on the extensive body of knowledge developed by Ray Rist and the members of INTEVAL to better understand how evaluation systems are built, how they function, and what impacts they may have on the evaluation practice. By revisiting and synthesizing this extensive body of work, we aim to extract key lessons for the design and implementation of evaluation systems that are both effective and adaptable to various contexts.

**Setting:** As researchers and evaluators working within academic institutions, our work is informed by interdisciplinary and international perspectives that we use to examine the development of evaluation systems across diverse political and administrative contexts.

**Intervention:** Not applicable.

**Research Design:** Not applicable.

**Data Collection and Analysis:** This article is based on a literature review focusing on the institutionalization of evaluation and the development of evaluation systems. Particular attention is given to the Comparative Policy Evaluation series established by Ray Rist, which offers an interdisciplinary and internationally comparative body of work spanning over three decades. Drawing on this series, we examined how evaluation systems have been conceptualized, implemented, and critiqued over time and across diverse national settings.

**Findings:** This article highlights that evaluation systems are shaped by contextual factors, institutional arrangements, and the contributions of key individuals. The effectiveness of these evaluation systems depends not only on organizational design but also on the capacity and engagement of evaluators, commissioners, and decision makers. The effective functioning of such systems fosters an environment favorable to evaluation, achieves a balance between supply and demand, and safeguards the independence and integrity of evaluation.

**Keywords:** *Evaluation Systems; Institutionalization of Evaluation; Public Sector Governance*

---

“The best evaluator in the world cannot make a difference in a context that does not favor evaluation.” This statement summarizes the discussions that both authors of this article had when sharing our preliminary thoughts about the influence of Ray Rist and the International Research Group for Policy and Program Evaluation (INTEVAL) on the institutionalization of evaluation around the world.

A considerable number of distinguished evaluators have advocated for the integration of evaluation into the regular operations of public organizations. At the outset, decision makers are typically reluctant to provide the public with information that might reflect shortcomings or deficiencies in the design or implementation of public programs (Jacob, 2023). The primary objective of an evaluation system is then to establish evaluation as a standard practice within public organizations:

The term “evaluation systems” refers to the procedural, institutional and policy arrangements shaping the evaluation function and its relationship to its internal and external environment—the aid agency, Parliament, other members’ evaluation departments, etc. This includes the evaluation function’s independence, the resources it relies upon and, not least, cultural attitudes to evaluation. (OECD, 2004, p. 4).

After we present what an evaluation system is, in the next section we analyze such systems’ pros and cons. We follow this with another section on the major contributions of the INTEVAL production under the influence and leadership of Ray Rist, identifying the key elements and pathways to establishing and implementing an evaluation system. Finally, we conclude that evaluation systems are shaped by contextual factors, institutional arrangements, and the contributions of key individuals, so that their effectiveness depends not only on organizational design but also on the capacity and commitment of evaluators, commissioners, and decision makers.

## Pros and Cons of an Evaluation System

The establishment of a national evaluation system necessitates a significant investment of time and energy. While such a system can enhance the effectiveness of public programs and policies, it also presents certain challenges that must be carefully considered. This section provides an overview of

the potential benefits and drawbacks associated with implementing an evaluation system.

One of the key benefits of a structured evaluation system is its ability to improve program management through better synchronization with organizational priorities. Integrating evaluation planning with organizational strategy ensures that evaluations align with key goals and decision-making needs (OECD, 2016). A national evaluation system can also help to balance the supply of and demand for evaluation (Palenberg & Paulson, 2021), while enhancing transparency by making findings more accessible to the public and supporting accountability for results.

Furthermore, the creation of an evaluation system raises expectations for evaluation (Chen, 2001) and strengthens individual and organizational competencies that support a robust evaluation infrastructure for effective data management and analysis.

Beyond the scope of individual programs or organizations, evaluation systems also influence the broader evaluation culture by promoting a learning-oriented environment and making evaluation more integral to decision-making (Leeuw & Furubo, 2008). As the OECD indicates, “Evaluation systems influence the demand for evaluation as well as its use, including particularly the dissemination, feedback and integration of evaluation findings” (OECD, 2004, p. 4).

Although a national evaluation system offers numerous advantages, it can also generate resistance, opposition, or unintended side effects.

The integration of a national framework often requires substantial efforts in terms of organizational structure and planning (Raimondo, 2018). It also requires adjustments to existing data systems or the development of entirely new infrastructures to meet reporting and evaluation requirements (Chen, 2001).

Some authors argue that without adequate resources and planning, evaluation systems risk becoming an administrative burden (Raimondo & Leeuw, 2022), compromising their independence, adaptability and relevance. This “evaluation capture” occurs when evaluators are inundated with “protocols, guidelines, standards, norms, criteria, templates, oversight, and reviews from national and supranational organizations, governments, NGOs, and agencies” (Leeuw & Pleger, 2023, p. 46). The resulting complexity may lead to overly rigid structures in which compliance takes precedence over meaningful inquiry.

A final challenge is balancing the institutionalization of evaluation practices with the risk of overstandardization. Common procedures can enhance consistency, but excessive uniformity

may limit adaptability across policy contexts. When evaluations become routine exercises bound by strict templates, they risk losing the capacity to produce nuanced and context-specific insights (Andersen, 2021).

In summary, balancing the benefits and challenges of national evaluation systems requires careful design and consideration.

## The Contribution of Ray Rist and INTEVAL to the Promotion and Understanding of Evaluation Systems

After considering the pros and cons, we present here the main characteristics of an “effective evaluation system.” To be effective, evaluation systems should be strategically structured to achieve several key objectives. First, they must enable evaluators, commissioners, and users to perform at their best by fostering a supportive and accountable environment. Second, they should balance supply and demand in a way that preserves the core mission and integrity of evaluation. Third, they need to guard against evaluation capture, ensuring that evaluation remains an independent and meaningful process rather than a bureaucratic exercise. Fourth, while standardization can enhance consistency, it should not be at the expense of methodological flexibility and contextual responsiveness. Finally, evaluation systems should not only strengthen the use and influence of evaluation results but also promote broader participation, collaboration, and democratic values in the evaluation process.

For decades, Ray Rist and INTEVAL have played a central role in shaping this conversation on institutionalization and evaluation systems. Ray Rist has coauthored with Jody Zall Kusek *Ten Steps to a Results-Based Monitoring and Evaluation System*, which is a reference for governments or organizations aiming to develop their evaluation system. INTEVAL members have published several books containing illustrations and advice on how to develop an evaluation system and have been seminal in the comparative study of the evaluation function and how it is institutionalized across different countries, organizations and systems.

From the first volume of the Comparative Policy Analysis series, the 1990 book *Program Evaluation and the Management of Government: Patterns and Prospects Across Eight Nations* under the leadership of Ray Rist, INTEVAL produced comparative studies on evaluation policy, and the way in which the evaluation function was institutionalized in different countries and

organizations. In that first book, Rist made the distinction between the management *of* evaluations and management *by* evaluations, highlighting that “the managerial input into the evaluation effort has considerable influence on the evaluation outcomes” (Leeuw & Sonnichsen, 1994, p. 8). In fact, the chapter in that book by Derlien (1990) states that the evaluation function and the conducting of evaluations should be institutionalized, as if that is not the case, their occurrence and subsequent use tend to be random.

Although several of the early books of the series have not directly dealt with institutionalization and evaluation systems per se, they set the scene for their study, helping to understand the practice of evaluation and the governmental and organizational arrangements around in different countries. This was the case with policy instruments (Bemelmans-Videc et al., 1998); the differences among budgeting, auditing, and evaluation (Gray et al., 1993) or performance auditing (Mayne & Zapico-Goñi, 1997); intergovernmental evaluation practices (Rieper & Toulemonde, 1997); and how to set effective evaluation capacity building in public organizations (Boyle & Lemaire, 1999).

However, in Volume 3, the 1994 book edited by Leeuw, Rist, and Sonnichsen *Can Governments Learn? Comparative Perspectives on Evaluation and Organizational Learning*, there is a specific attention to the theoretical understandings of organizational learning processes; that is, how organizations assimilate and use information, recognizing that the evaluation function must be, among other aspects, internally interlinked with organizational aspects (in what later will be recognized as evaluation systems). Although this volume is rather focused on evaluation as an aid to organizational learning, there is an already clear understanding that evaluations, to be useful, should be organizationally integrated and the evaluation function must be constructively managed. Moreover, as recognized by its editors, the different contributions of the book demonstrated that institutionalization is important irrespective of whether evaluations are performed internally or externally (Leeuw & Sonnichsen, 1994).

A seminal book in the study of evaluation institutionalization and evaluation systems was Volume 9 of the series, *International Atlas of Evaluation*, edited by Furubo, Rist, and Sandahl (2002). It was a significant effort to present and analyze empirical cases in a comparative way, describing evaluation systems in 21 countries and development aid agencies, the European Union, and the World Bank. The idea of the book was to

describe how evaluation activities were organized in different countries, and to look for global trends and developments, as well as to explain them. The countries were ranked following nine “indicators”<sup>1</sup> which set the first analytical framework for studying evaluation systems and the institutionalization of the evaluation function in each country or organization. This analytical framework became the basis for any study on evaluation institutionalization and has been widely used, along with the later review by Jacob, Speer, and Furubo (2015); the works by Varone and Jacob (2004); and the later works by Stockmann et al. (2020, 2022, 2023). Any new analytical effort is based on that first work done by INTEVAL. Even the most innovative aspect of the most recent analytical framework proposed by Stockmann et al., which is developing, along with a political system and a system of professionalization, a social subsystem (use of evaluations by civil society, public discourse and perception of evaluation, and demand of evaluation by civil society), is based on the former indicator related to the presence of a national discourse concerning evaluation.

In Volume 12 of the series, *From Studies to Streams: Managing Evaluative Systems*, edited by Rist and Stame (2006), the authors used the metaphor of streams to identify the different streams of evaluative knowledge that can take place in an evaluation system. This use of the term “streams” instead of “studies” indicates a deliberate intention to go beyond the definition of the “evaluation work as a production of individual, discrete studies for individual, discrete users” (Rist, 2006, p. 284). This narrow conception of individual studies leads to “evaluation amnesia” and does not take advantage of a proper evaluation system, which allows organizations to learn more quickly, systematically, and comprehensively, by creating, mediating, and distributing new knowledge. This evolution from studies to streams is thus well connected to evaluation systems. As Martinaitis and colleagues stated several years later, “The characteristics of individual evaluations cannot explain why some institutions, regions and countries tend to systematically produce higher quality evaluations, which are better integrated into the decision-making process” (Martinaitis et al., 2019, p. 47).

Other books in the series do not directly address the issue of evaluation institutionalization and systems but certainly contribute to a better understanding of it by shedding some light on issues that are important components or factors to consider when studying the institutionalization of evaluation. These are specifically the cases of Volume 13, on the role of evaluative information and its use in the public arena (Boyle et al., 2007); Volume 18, where complexity and its understanding in evaluation practices and systems are in focus (Forss et al., 2011); and Volume 19, on the diversity of evaluation cultures that should necessarily be considered in an evaluation system (Barbier & Hawkins, 2012).

In the most recent volumes, a less optimistic, or at least a more critical, view about the institutionalization of evaluation is presented. This is parallel to the evolution of evaluation, its own institutionalization and “industrialization.” *The Evaluation Enterprise: A Critical View* (Furubo & Stame, 2018) tackled some consequences of the overinstitutionalization of evaluation, raising a “skeptical turn in evaluation” (Dahler-Larsen, 2018), and of a professionalization of evaluators not necessarily favoring their action in the public interest (Schwandt, 2018). Palenberg and Paulson (2021) in *The Realpolitik of Evaluation: Why Demand and Supply Rarely Intersect*, identify some gaps between the demand for evaluation and the supply of information provided by evaluators, something that any evaluation system should avoid, and present some attempts to bridge this gap through the use of third parties, increased stakeholder involvement, and the incorporation of social science models to strengthen the theories of change. Finally, in the volume *Changing Bureaucracies: Adapting to Uncertainty, and How Evaluation Can Help* (Perrin & Tyrrell, 2021), a discussion on how public sector bureaucracies need to adapt to the reality of unprecedented levels of uncertainty and complexity, there is also a critical perspective on how evaluation systems can fall in bureaucratic capture, identifying “a number of pernicious facets that can, and sometimes do, accompany the institutionalization of evaluation within bureaucracies” (Raimondo & Leeuw, 2022, p. 143) The authors describe three dysfunctions evaluation systems suffer in bureaucracies: (1) Irrationality of rationalization, “when

<sup>1</sup> These were: 1) Evaluation takes place in many policy domains; 2) Supply of domestic evaluators in different disciplines; 3) National discourse concerning evaluation; 4) Professional organizations; 5) Degree of institutionalization–Government; 6) Degree of institutionalization–Parliament; 7) Pluralism of

institutions or evaluators performing evaluations within each policy domain; 8) Evaluation within the Supreme Audit Institution; and 9) Proportion of outcome evaluations in relation to output and process evaluations (Furubo et al., 2002, pp. 7–10).

bureaucracies adapt their missions to fit the existing rules of the game” (p. 151); (2) Cultural contestation, “when the various constituencies of an organization clash over competing perspectives of the organization’s mission and performance” (p. 152); and (3) Bureaucratic universalism, already described by Barnett & Finnemore, 2004, when bureaucracies become obsessed with their own rules at the expense of their primary missions. But they also propose five ways for breaking out of the bureaucratic capture: (1) Adhere to a “skeptical turn” (Dahler-Larsen, 2018); (2) Scale back results-based management principles; (3) Cut down on general rules; (4) Rely more on judgment and less on indicators; and (5) help evaluators change their behavior in order to be more resilient (Leeuw & Pleger, 2023).

Over the past 40 years, the INTEVAL book series has consistently occupied a leading position in the field by critically examining the strengths and limitations of institutionalization processes and evaluation systems

## Conclusion

By taking stock of numerous studies and decades of experience across diverse national contexts, scholars have identified the core features of effective evaluation systems. These systems succeed when they align institutional frameworks with the contributions of skilled individuals, ensuring that evaluation remains both rigorous and adaptable to the complexities of decision-making and governance. Effective evaluation systems institutionalize evaluation practices and create an environment in which competent professionals can make meaningful contributions. At the heart of such systems is a dynamic interaction between institutional arrangements and individual agency. While organizations establish the structural foundations for evaluation to take place, its effectiveness ultimately depends on the evaluators, commissioners, and decision makers who shape both the supply of and demand for evaluation knowledge.

Despite the common understanding of the elements that contribute to the effectiveness of evaluation systems, their implementation varies significantly across countries and organizations. Recent research on the institutionalization of evaluation highlights the absence of a single explanatory model for why some systems are more fully developed than others (Stockmann et al., 2020, 2022, 2023). Instead, these systems reflect the distinct political, organizational, and

professional interests of the actors who have designed and promoted them (Dediu et al., 2019).

Understanding these variations calls for continued research to build on the foundational work initiated by Ray Rist and INTEVAL, which continues to shape the global conversation on evaluation systems.

## References

- Andersen, N. A. (2021). The power and paradoxes of evaluation systems – increasing use but impeding change. *Scandinavian Journal of Public Administration*, 25(3/4), 39–59. <https://doi.org/10.58235/sjpa.v25i3-4.7063>
- Barbier J.-C., & Hawkins, P. (Eds.). (2012). *Evaluation cultures: Sense-making in complex times* (Comparative Policy Evaluation Vol. 19). Transaction Publishers.
- Barnett, M. N., & Finnemore, M. (2004). *Rules for the world: International organizations in world politics*. Cornell University Press.
- Bemelmans-Videc, M.-L., Rist, R. C., & Vedung, E. (Eds.). (1998). *Carrots, sticks and sermons: Policy instruments and their evaluation* (Comparative Policy Evaluation Vol. 7). Transaction Publishers.
- Boyle, R., & Lemaire, D. (Eds.). (1999). *Building effective evaluation capacity: Lessons from practice* (Comparative Policy Evaluation Vol. 8). Transaction Publishers.
- Boyle, R., Breul, J., & Dahler-Larsen, P. (Eds.). (2007). *Open to the public: Evaluation in the public arena* (Comparative Policy Evaluation Vol. 13). Transaction Publishers.
- Chen, H.-T. (2001). Development of a national evaluation system to evaluate CDC-funded health department HIV prevention programs. *American Journal of Evaluation*, 22(1), 55–70. <https://doi.org/10.1177/109821400102200106>
- Dahler-Larsen, P. (2018). The skeptical turn in evaluation. In J.-E. Furubo & N. Stame (Eds.), *The evaluation enterprise: A critical view* (Comparative Policy Evaluation). Routledge.
- Derlien, H. U. (1990). Genesis and structure of evaluation efforts in comparative perspective. In R. C. Rist (Ed.), *Program evaluation and the management of government: Patterns and prospects across eight nations*. Transaction Publishers.
- Dediu, B. I., Moga, L. M., & Mirica, S. C. (2019). An overview of the trends in the evaluation system for the public management performance. In M. Bilgin, H. Danis, E. Demir, & U. Can (Eds.), *Eurasian economic perspectives: Proceedings*

- of the 24<sup>th</sup> Eurasia Business and Economics Society Conference (pp. 95-103). Springer.
- Forss, K., Marra, M., & Schwartz, R. (Eds.). (2011). *Evaluating the complex: Attribution, contribution and beyond* (Comparative Policy Evaluation Vol. 19). Transaction Publications.
- Furubo, J.-E., Rist, R. C., & Sandahl, R. (Eds.). (2002). *International atlas of evaluation* (Comparative Policy Evaluation Vol. 9). Transaction Publications.
- Furubo, J.-E., & Stame, N. (Eds.). (2018). *The evaluation enterprise: A critical view*. Routledge.
- Gray, A. G., Jenkins, W., I., & Segsworth, R. V. (Eds.). (1993). *Budgeting, auditing and evaluation: Patterns and integration in seven governments* (Comparative Policy Evaluation Vol. 2). Transaction Publications.
- Jacob, S. (2023). The institutionalization of evaluation around the globe: Understanding the main drivers and effects over the past decades. In F. Varone, S. Jacob, & P. Bundi (Eds.), *Handbook of public policy evaluation* (pp. 187–205). Edward Elgar. <https://doi.org/10.4337/9781800884892.00021>
- Jacob, S., Speer, S., & Furubo, J.-E. (2015). The institutionalization of evaluation matters: Updating the International Atlas of Evaluation 10 years later. *Evaluation*, 21(1), 6–31. <https://doi.org/10.1177/1356389014564248>
- Leeuw, F. L., & Sonnichsen, R. C. (1994). Introduction. In F. L. Leeuw, R. C. Rist, & R. C. Sonnichsen (Eds.), *Can governments learn?: Comparative perspectives on evaluation and organizational learning*. Transaction Publishers.
- Leeuw, F. L., & Furubo, J.-E. (2008). Evaluation systems: What are they and why study them? *Evaluation*, 14(2), 157–169. <https://doi.org/10.1177/1356389007087537>
- Leeuw, F., & Pleger, L. (2023). Evaluation capture, evaluator resilience, and the need for competencies of evaluators. *Journal of MultiDisciplinary Evaluation*, 19(46), 45–53. <https://doi.org/10.56645/jmde.v19i46.863>
- Martinaitis, Ž., Christenko, A., & Kraučūnienė, L. (2019). Evaluation systems: How do they frame, generate and use evidence? *Evaluation*, 25(1), 46–61. <https://doi.org/10.1177/1356389018802135>
- Mayne, J., & Zapico-Goñi (Eds.). (1997). *Monitoring performance in the public sector: Future directions from international experience* (Comparative Policy Evaluation Vol. 5).
- OECD. (2004). *Evaluation systems in DAC members' agencies: A study based on DAC Peer Reviews*.
- OECD. (2016). *Evaluation systems in development co-operation: 2016 review*. <https://doi.org/10.1787/9789264262065-en>.
- Palenberg, M., & Paulson, A. (Eds.). (2021). *The realpolitik of evaluation: Why demand and supply rarely intersect* (Comparative Policy Evaluation). Routledge.
- Perrin, B., & Tyrell, T. (Eds.). (2021). *Changing bureaucracies: Adapting to uncertainty, and how evaluation can help* (Comparative Policy Evaluation). Routledge.
- Raimondo, E. (2018). The power and dysfunction of evaluation systems in international organizations. *Evaluation*, 24(1), 26–41. <https://doi.org/10.1177/1356389017749068>
- Raimondo, E., & Leeuw, F. L. (2022). Evaluation systems and bureaucratic capture: Locked. In B. Perrin & T. Tyrell (Eds.), *Changing bureaucracies: Adapting to uncertainty, and how evaluation can help* (pp. 143–160). Routledge.
- Rieper, O., & Toulemonde, J. (Eds.). (1997). *Politics and practices of intergovernmental evaluation* (Comparative Policy Evaluation Vol. 4). Transaction Publishers.
- Rist, R. C. (1990). Managing of evaluations or managing by evaluations: Choices and consequences. In R. C. Rist (Ed.), *Program evaluation and the management of government: Patterns and prospects across eight nations* (Comparative Policy Evaluation Vol. 1). Transaction Publishers.
- Rist R. C. (2006). Conclusion. A Brief Critique. In Rist R. C., & Stame, N. (Eds.). (2006). *From studies to streams: Managing evaluative systems* (Comparative Policy Evaluation Vol. 12). Transaction Publishers.
- Rist R. C., & Stame, N. (Eds.). (2006). *From studies to streams: Managing evaluative systems* (Comparative Policy Evaluation Vol. 12). Transaction Publishers.
- Schwandt, T. A. (2018). The concerted effort to professionalize evaluation practice: Whither are we bound? In J.-E. Furubo & N. Stame (Eds.), *The evaluation enterprise: a critical view*. Routledge.
- Stockmann, R., Meyer, W., & Taube, L. (Eds.). (2020). *The institutionalisation of evaluation in Europe*. Palgrave Macmillan.
- Stockmann, R., Meyer, W., & Szentmarjay, L. (Eds.). (2022). *The institutionalisation of evaluation in the Americas*. Palgrave Macmillan.

- Stockmann, R., Meyer, W., & Zierke, N. (Eds.). (2023). *The institutionalisation of evaluation in Asia-Pacific*. Palgrave Macmillan.
- Stockmann, R., Meyer, W., & Zierke, N. (Eds.). *The institutionalisation of evaluation in Asia-Pacific*. Palgrave Macmillan.
- Varone, F., & Jacob, S. (2004). Institutionnalisation de l'évaluation et nouvelle gestion publique: Un état des lieux comparatif. *Revue internationale de politique comparée*, 11(2), 271–292. <https://doi.org/10.3917/ripc.112.0271>