



# The Institutionalisation of Evaluation in the Americas

*Edited by* Reinhard Stockmann  
Wolfgang Meyer · Laszlo Szentmarjay

palgrave  
macmillan

# The Institutionalisation of Evaluation in the Americas

Reinhard Stockmann · Wolfgang Meyer ·  
Laszlo Szentmarjay  
Editors

# The Institutionalisation of Evaluation in the Americas

palgrave  
macmillan

*Editors*

Reinhard Stockmann  
Saarland University  
Saarbrücken, Germany

Wolfgang Meyer  
Saarland University  
Saarbrücken, Germany

Laszlo Szentmarjay  
Center for Evaluation (CEval)  
Saarbrücken, Germany

ISBN 978-3-030-81138-9      ISBN 978-3-030-81139-6 (eBook)  
<https://doi.org/10.1007/978-3-030-81139-6>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer  
Nature Switzerland AG 2022

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover image: © Ashway/Alamy Stock Photo

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

## PREFACE AND ACKNOWLEDGEMENTS

This book is the second one in the Studies in the Institutionalization of Evaluation-series with the goal of examining the institutionalisation of evaluation viewed from a global perspective. Two further publications on Asia-Pacific as well as Africa will follow.

After the publication of the volume on Europe<sup>1</sup> last year, this book provides an overview of the institutionalisation of evaluation in the countries of North, Central and South America. This allows comparisons to be made not only between individual countries, but also with Europe.

Despite all the political, economic, social and cultural differences, it becomes evident that evaluation is still a young discipline that is most likely manifested in politics and administration, but little in society and science. Although nearly every country in the Americas (as well as in Europe) has Voluntary Organisations for Professional Evaluation (VOPEs), evaluation lacks a strong social and scientific lobby.

This is well exemplified by the COVID 19-pandemic currently spreading in all countries of the world. Although evaluation should be more in demand than ever, in many countries it is not used to assess the success of hygiene concepts and measures. Such evaluations could not only identify the best concepts and measures and reduce the incidence of

<sup>1</sup> Stockmann, R. Meyer, W., & Taube, L. (2020). *The Institutionalisation of Evaluation in Europe*. Cham/Switzerland: Springer Nature Switzerland/Palgrave Macmillan.

infection by implementing them throughout the country, but the politically prescribed measures would also gain legitimacy and thus possibly more acceptance among the population.

A second problem is that although many studies are conducted by physicians, epidemiologists, statisticians, economists and sociologists and arrive at valuable and interesting findings, they are not perceived as a contribution of evaluation to scientific knowledge and to overcoming the crisis.

This is a problem for a young profession that still needs to establish itself more firmly in politics, society and science if it is to survive in the long term. To increase its dissemination and recognition, evaluation must prove its usefulness.

The urgent need for this is illustrated by the analysis of the institutionalisation of evaluation in the Americas and no less in Europe—as already shown in the previous volume.

If the institutionalisation of evaluation and thus its dissemination and development is to be promoted, then an inventory of the present situation is the first step towards this. This is the only way to identify the driving and hindering forces, where there is a need for change and how evaluation needs to position itself in politics, society and science so that it can unfold its useful potential. Who should know this better than us—evaluators ourselves?

We hope that this book is a guide to contribute constructively to these themes.

First of all, in order for such a book to be produced at all, we needed knowledgeable country experts. Therefore, our biggest thanks go to the authors of the case studies, who have taken on the Herculean task of analysing evaluation with its diversity and complexity across all sectors in politics, society and science with remarkable motivation, meticulousness and prudence. They deserve thanks not only for sharing their knowledge with us, but also for adhering to the restrictive guidelines of an analytical grid that enabled us to compare the countries with each other and with the results of the European volume.

If this is to be crowned by success with 35 authors working on it, a strong coordinating force is needed to ensure that the methodological, content-related and time-related guidelines are adhered to. Laszlo Szentmarjay, Felipe Ramirez-Kaiser and Lena Taube have accomplished this task, which should not be underestimated. Therefore, they deserve our special thanks.

As the majority of the case studies had been written in Spanish language there was a special need translating these into English. A further thanks goes therefore to Daniela Mora Vera from Ecuador, who took on this task.

Last but not least, every book publication naturally requires the support of hard-working staff who transform the text into a print-ready version. Here we would particularly like to thank Angelika Nentwig, Nicole Ebel, Myriel Mohra and Zahida Moulay.

We hope that the second volume of this series will also contribute to understanding the development of evaluation so far and to drive institutionalisation of evaluation forward in state, society and science.

Büstadt/Saarbrücken,  
Germany  
December 2020

Reinhard Stockmann  
Wolfgang Meyer  
Laszlo Szentmarjay

# CONTENTS

## Part I Introduction

- 1 The Institutionalisation of Evaluation: Theoretical Background, Analytical Concept and Methods** 3  
Wolfgang Meyer, Reinhard Stockmann,  
and Laszlo Szentmarjay

## Part II National Developments

- 2 Evaluation in Argentina** 41  
Pablo Rodríguez-Bilella and Esteban Tapella
- 3 Evaluation in Bolivia** 65  
Silvia Tatiana Salinas Mulder, Martha Lanza Meneses,  
Scarleth Flores Calle, María Dolores Castro Mantilla,  
and Paul Mauricio Villarroel Via
- 4 Evaluation in Brazil** 93  
Verena Jacques Dolabella
- 5 Evaluation in Canada** 143  
Benoît Gauthier, Robert E. Lahey, and Steve Jacob
- 6 Evaluation in Chile** 171  
Claudia Olavarría Manríquez and Andrea Peroni Fiscarelli



<b>7</b>	<b>Evaluation in Colombia</b>	<b>201</b>
	Maria Gladys Álvarez Basabe, María Elena Manjarrés De Mendoza, Luz Stella Uricoechea Morales, María Paula Uricoechea Castellanos, and Belkis Esperanza Vergara Pérez	
<b>8</b>	<b>Evaluation in Costa Rica</b>	<b>239</b>
	Mayela Cubillo Mora and María Camila León Betancourth	
<b>9</b>	<b>Evaluation in Ecuador</b>	<b>267</b>
	Cinthia Josette Arévalo Gross and Lourdes Cumandá Montesdeoca Espín	
<b>10</b>	<b>Evaluation in Mexico</b>	<b>297</b>
	Janett Salvador Martínez and Jaqueline Meza Urías	
<b>11</b>	<b>Evaluation in Peru</b>	<b>323</b>
	Brenda Bucheli del Águila, Susana Guevara Salas, and Emma Lucía Rotondo Dall’Orso	
<b>12</b>	<b>Evaluation in the United States of America</b>	<b>355</b>
	Scott I. Donaldson, Stewart I. Donaldson, and Jessica A. Renger	
<b>Part III Transnational Organisations and Networks</b>		
<b>13</b>	<b>CLEAR LAC—Centers for Learning on Evaluation and Results—Latin America and the Caribbean</b>	<b>381</b>
	Claudia Maldonado Trujillo	
<b>14</b>	<b>Contributions of the Inter-American Development Bank Group (IDB Group) to the Institutionalisation of Evaluation in Latin America and the Caribbean</b>	<b>403</b>
	Ana Maria Linares, Anna Funaro Mortara, and Melanie Putic	
<b>15</b>	<b>The Independent Evaluation Group (IEG) of the World Bank Group: Influences on Evaluation Structures and Practices Globally and in the Americas</b>	<b>417</b>
	Maurya West Meiers	

**Part IV Synthesis**

- 16 The Institutionalisation of Evaluation in the Americas:  
A Synthesis** 451  
Reinhard Stockmann and Wolfgang Meyer
- 17 The Institutionalisation of Evaluation in Europe  
and the Americas: A Comparison** 509  
Reinhard Stockmann and Wolfgang Meyer

## NOTES ON CONTRIBUTORS

**Maria Gladys Álvarez Basabe, Ph.D.** in educational innovation; university teacher, with 40 years of experience in teacher training. MEN-Colombia, technical advisor in programmes for diversified secondary education, educational promotion in rural areas and the evaluation division. Coordinator of research seedbeds in undergraduate programmes. Consultant for the Ministerio del Interior for the Community Integration and Development directorate.

**Cinthia Josette Arévalo Gross** is an evaluation specialist at the Office of Evaluation and Oversight of the IDB. She has a Ph.D. in Public Policy and Administration, and master's degrees in Development Economics and Public Policy. Josette has worked as the Executive Director of the National Institute of Educational Assessment of Ecuador and has M&E work experience in academia, the public sector and international organisations.

**Brenda Bucheli del Águila** Social Psychologist with a Master's in Human Resources and Knowledge Management, specialising in monitoring and evaluation, organisational development, information management. More than 20 years of national and international experience, with various sectors and types of institutions, in technical and managerial roles. President of EvalPeru (2017–2019). Postgraduate professor since 2006. Director of USAID's Monitoring, Evaluation, and Learning for Sustainability Project (2019–2020).

**María Dolores Castro Mantilla** Anthropologist. Professional performance in research, systematisation, and evaluation of projects and programmes in different areas of development mainly: gender, intercultural relations, sexual and reproductive rights. Member of academic councils, research institutes, evaluation networks. Professor and Coordinator of the Observatory of Maternal and Neonatal Mortality of the Universidad Mayor de San Andrés.

**Mayela Cubillo Mora** Former Vice-Dean of the Faculty of Economic Sciences, University of Costa Rica.

Former Director of the School of Public Administration, University of Costa Rica as well as former Director of the Center for Research and Training in Public Administration, University of Costa Rica.

She has taught Master's courses at the Universidad Autónoma de Honduras, at the Universidad de San Carlos in Guatemala, at the Universidad Autónoma de República Dominicana.

Mayela does research in the areas of evaluation of social public policies, in social economy and local government issues.

Author of four books and a large number of articles.

**Verena Jacques Dolabella** is an Independent Consultant on M&E, Collaborative Impact associate and currently researcher at CLEAR LAB. Her main interests are supporting public and development organisations in Brazil and Latin America using participatory approaches in the areas of education, health, women's empowerment and advocacy. She is also a member of the Brazilian M&E Network.

**Scott I. Donaldson** is a Postdoctoral Scholar in evaluation, statistics and measurement at the University of California, San Diego School of Medicine. He received his Ph.D. in Psychology with a concentration in evaluation and applied research methods. His research focuses on the design and evaluation of health behaviour interventions in the workplace.

**Stewart I. Donaldson** is a distinguished university professor, the executive director of the Claremont Evaluation Center, and director of The Evaluators' Institute at Claremont Graduate University. He is deeply committed to improving lives through research, evaluation and education. He has published extensively on evaluation theory and practice, received numerous career achievement awards for his evaluation work and served as President of the American Evaluation Association in 2015.

**Scarleth Flores Calle** Industrial Engineer; Master's in Local Development, M.B.A. in Strategic Leadership and Master's in Business Administration. Scarleth has worked in organisational strengthening, monitoring and evaluation in areas of local economic development and health, mainly linked to the rural and forestry sectors. She has worked also for international and state organisations in Bolivia and Honduras.

**Benoît Gauthier** a Credentialed Evaluator, Benoît Gauthier is dedicated to the Canadian Evaluation Society (CES) and to associative life in the evaluation world in general. He is a Consultant in evaluation, a teacher on evaluation, a volunteer researcher on evaluation practice in Canada, and the co-editor of a textbook on social research methods.

**Susana Guevara Salas** Sociologist, with a Master's in Evaluation of Public Policies and Social Management, with more than 20 years of work in the development of evaluations, monitoring systems and systematisations, using mixed evaluation methodologies. She has work experience in different Latin American countries. Vice Coordinator of ReLAC (2015–2017).

**Steve Jacob** is director of PerfEval, a research laboratory on public policy performance and evaluation. Steve Jacob conducts research into the mechanisms of performance management and evaluation. He authored or edited ten books and numerous journal articles. Steve Jacob serves as an expert in public administration reform and holds a research chair on public administration in the digital era.

**Robert E. Lahey** is the founding head Canada's Centre of Excellence for Evaluation; former Evaluation Head in four Canadian agencies. Since 2004, international advisor on evaluation capacity building, use and results-oriented monitoring and evaluation, including World Bank, regional banks, UN agencies, directly with many countries globally. Extensive writings and frequent invited speaker to global/regional events.

**Martha Lanza Meneses** Master in Social Policy Administration (FLACSO) has held management positions in public and private bodies. She has developed advocacy processes in national and municipal public policies, and has managed development projects. She is a Consultant, Evaluator and Activist for gender equality. Member of REDWIM, REDMEBOL, AWID and other networks.

**María Camila León Betancourth** Political scientist with concentration on conflict resolution and peace studies from the Pontificia Universidad Javeriana (Bogota, COL); training specialist on human rights and international humanitarian law at the Universidad Nacional de Colombia; passionate on human rights issues, international law, conflict resolutions, migration and refugees.

**Ana Maria Linares** is a Senior Advisor in the Office of Evaluation and Oversight (OVE) of the Inter-American Development Bank (IDB) Group. She leads the corporate evaluation cluster and is also responsible for implementing OVE's evaluation capacity development strategy in Latin America and the Caribbean. She has worked more than 20 years in various capacities across the IDB Group, including in the Office of Strategic Planning and Development Effectiveness; the Environment and Infrastructure Department covering Central America and the Caribbean; and the Legal Department. She holds a Master of Laws from Georgetown University; a Master of Arts from the Johns Hopkins School of Advanced International Studies in Washington, DC; and a Juris Doctor from Universidad de los Andes in Bogota, Colombia.

**Claudia Maldonado Trujillo** is an academic based in Mexico-city (Center for Research and Teaching in Economics, CIDE). She specialises in social policy in Latin America, programme evaluation and capacity building in M&E. She was the founding Director of the CLEAR Center for Latin America and the Caribbean (2011–2016) and currently serves as academic researcher at the National Council for the Evaluation of Social Development Policy (CONEVAL) in Mexico. She is a member of the Evaluation Advisory Panel at UNDP and IPDET's Advisory Board.

Among her publications are 'Two Persistent Dimensions of Democracy', with Michael Coppedge and Ángel Álvarez (*Journal of Politics*, 2008); 'Enhancing Accountability Through Results-Oriented Monitoring and Evaluation Systems', with Sonia Ospina and Nuria Cunil (*Handbook of Latin American Public Administration*, 2019), 'La integración de políticas públicas para el desarrollo: Brasil y México en perspectiva comparada' with Mariana Magaldi (eds.) (*CIDE* 2014). With Gabriela Pérez Yarahuán, she has co-edited several volumes on the origins of evaluation as a discipline, impact evaluation and M&E systems in Latin America (CLEAR 2016, 2017, 2020).

Claudia Maldonado has extensive teaching experience on programme evaluation and public policy in the global south.

**María Elena Manjarrés De Mendoza, Ph.D.** in Development Studies, co-author of *Research as a Pedagogical Strategy* supported by ICT, with experience in public procurement and in the relationship and management with government entities and private companies oriented to obtain strategic associations and in the structuring, execution, systematisation and evaluation of projects, programmes and policies of education and training of teachers and in the development of software, digital content and physical and virtual pedagogical materials for these processes.

**Wolfgang Meyer** is an Associate Professor at Saarland University, Saarbrücken, and an adjunct professor at the Ugandan Technology and Management University, Kampala; Vice-Director of the Center for Evaluation (CEval); a founding member of the German Evaluation Society (DeGEval). A sociologist with a focus on Empirical Methodology, Wolfgang is an evaluation specialist in the fields of environment, labour market, regional development and qualitative and quantitative methods. The projects he realised were involved in local companies but also in organisations all around the globe.

**Jaqueline Meza Urías** has worked as an Academic Researcher and public official for more than 15 years and as a Consultant for government for more than three years on issues related to evaluation of public programmes and public administration in general. Jaqueline has been member of the Mexican Association of Evaluators (ACEVAL) since 2016.

**Lourdes Cumandá Montesdeoca Espín** is an expert in applied econometrics, economic policy and quantitative economics in the Public Sector. She has worked in the most prominent economic research centres in Ecuador such as: INEC, CEF-SRI, SENPLADES, SIISE and the Central Bank. Now she is a University Professor and dictates Statistics, Econometrics and Quantitative Methods.

**Anna Funaro Mortara** is a Research Fellow in the Office of Evaluation and Oversight (OVE) of the Inter-American Development Bank (IDB) Group. While at OVE she has contributed to several corporate, country and project evaluations. Before joining OVE she was a Project Manager for J-PAL (Abdul Latif Jameel Poverty Action Lab) and worked in local governments in Brazil. She studied in Sao Paulo (Getulio Vargas Foundation—MPA and B.Sc. in Economics), Paris (Sciences-Po Paris and Université la Sorbonne-Paris IV) and Shanghai (Jiao Tong University of Shanghai).

**Claudia Olavarría Manríquez** Sociologist, Diploma and Master in Gender, Society and Politics. She is co-chair of EvalYouth, Consultant and Lecturer in planning and evaluation of social programmes and projects, having supported more than 30 public, private and third sector organisations in Chile and the region.

**Andrea Peroni Fiscarelli** Professor of History, Sociologist and Doctor in American Studies. She has been co-founder and coordinator of the Chilean Evaluation Network. Andrea is an academic at the University of Chile, where she coordinates the interdisciplinary nucleus of evaluation research. She works as a Consultant in social policy and its evaluation, in Chile and in the region. She is a member of the Executive Committee of ReLAC for two periods.

**Melanie Putic** is Technical Support Administrative Coordinator in the Office of Evaluation and Oversight (OVE) of the Inter-American Development Bank (IDB) Group. She has previously worked with the IDB Group's Independent Consultation and Investigation Mechanism (MICI) in Washington DC, the United Nations Department of Economic and Social Affairs (UNDESA) in New York NY, and the Red Cross in Alberta, Canada. She holds a master's degree in Environment and Sustainable development from the United Nations University for Peace in San Jose, Costa Rica, and a bachelor's degree in Sociology from Mount Royal University in Calgary, Canada.

**Jessica A. Renger** is a Ph.D. student in the Evaluation and Applied Research Methods programme at Claremont Graduate University. She has extensive evaluation experience both internally in collaboration with the Claremont Evaluation Center and externally as an independent Consultant. She has worked on projects nationally and internationally in the health, education and leadership sectors. She specialises in theory-driven programme evaluation, effective interpersonal communication and health promotion.

**Pablo Rodríguez-Bilella, Ph.D.** in Sociology from Sussex University (UK) and member of the CONICET—Argentine Research Council—as well as professor of Anthropology at Universidad Nacional de San Juan (Argentina). He was coordinator of the working group that developed the 'Evaluation Standards for Latin America' and he is member of the coordination team in charge of EvalParticipativa, a research programme



and community of practice about participative evaluation in Latin America and the Caribbean: <https://evalparticipativa.net/en/>.

**Emma Lucía Rotondo Dall’Orso** Social Anthropologist Advisor in Logotherapy with more than 30 years of experience in 17 countries of the Latin American region. She has directed the Regional Program for Capacity Building in Monitoring and Evaluation of PREVAL (2004–2015). He is a founding member of ReLAC and is currently a member of the Board of Directors of the Peruvian Evaluation Network EvalPeru. She is the regional coordinator of ReLAC’s community of practice of evaluators ‘M&E of human rights, conflict and violence programs’.

**Silvia Tatiana Salinas Mulder** Bolivian anthropologist, 30 years of professional experience. Advocate and facilitator of transformative evaluation approaches. Led and participated in complex international evaluations. Particularly interested in power issues, ethics, competencies, gender and intercultural relations. International speaker and author. 2016–2018 President of the Bolivian VOPE REDMEBOL, 2017–present Chair of ReLAC and since April 2020 President of IOCE.

**Janett Salvador Martínez** is an Evaluator and Consultant from Mexico. She has extensive experience in the social and economic evaluation of productive and social projects, as well as in monitoring and evaluation. Janett was founder member and First President of the Mexican Association of Evaluators (ACEVAL).

**Reinhard Stockmann** Professor for Sociology at Saarland University, Saarbrücken; Founder and Director of the Center for Evaluation (CEval); Managing Director of the German-speaking MA-Course Evaluation and of the English-speaking Blended Learning Master of Evaluation; Executive Editor of the German Zeitschrift für Evaluation (*Journal of Evaluation*); Founding member of the German Society for Evaluation (*DeGEval*). During his 50 years of evaluation experience, he conducted several hundred evaluation studies and has led numerous Evaluation Capacity Building projects in Africa, Asia and Latin America. He published about 50 books and about 300 articles dealing with the subject evaluation, quality development, development policy, development cooperation, vocational training, environment, sociology, etc. His textbooks have been translated into six languages.

**Laszlo Szentmarjay** (M.A.) is Research Associate at Centre for Evaluation (CEval) at Saarland University and project manager at CEval GmbH. He is responsible for the conception and implementation of evaluations and M&E systems in the field of development cooperation, environment and education. He is also doing research in the field of evaluation.

**Esteban Tapella** Magister in Development Studies and Ph.D. in Human Ecology. Professor and Researcher in the National University of San Juan, Argentina. Independent Consultant in evaluation and documentary photography. Member of EvaluAR (Argentine Evaluation Network) and RELAC (Latin American and Caribbean Network of Monitoring, Evaluation and Systematisation).

**María Paula Uricoechea Castellanos** Social anthropologist with emphasis in culture and power studies. Has worked among different subjects concerning economic and political anthropology, including fieldwork research with artists, bullfighters, gender studies and evaluation of public policies. The latest research involves the Venezuelan migration phenomenon in Colombia, working with ILO.

**Luz Stella Uricoechea Morales, Ph.D.** Honoris Causa in evaluation and quality. 35 years of experience as a teacher in basic, secondary and university education. Advises on curricular design, innovation projects in public and private institutions. Participant in research groups dedicated to the evaluation of public policies, innovation and educational inclusion. Experience in teacher training on issues related to curriculum design and management.

**Belkis Esperanza Vergara Pérez** Social Worker from Universidad de Cartagena—Colombia with Master in Education Candidate, experienced in educational coordination and social projects, university teaching and teacher trainer in public and private educational institutions.

**Paul Mauricio Villarroel Via** Economist with a Master Degree in Public Policy Evaluation. Specialist in planning, monitoring and evaluation of development programmes for national and international institutions. Associate Researcher at Fundación ARU in the Monitoring and Evaluation Group. His studies include impact evaluation, poverty, education and rural development.

**Maurya West Meiers** is a senior evaluation officer in the World Bank's Independent Evaluation Group. She has trained professionals at the International Program for Development Evaluation Training (IPDET) and across Africa, East Asia, South Asia, Europe and Latin America on monitoring and evaluation methods and systems. Maurya is co-author of the book *A Guide to Assessing Needs: Essential Tools for Collecting Information, Making Decisions, and Achieving Development Results*.

# LIST OF FIGURES

Fig. 1.1	Selected countries and transnational organisations for the case studies	23
Fig. 7.1	Typology of evaluations	209
Fig. 7.2	Sectors carrying out evaluations	209
Fig. 7.3	Evaluation dimensions and components	225
Fig. 9.1	Ecuadorian state functions and institutions	269
Fig. 13.1	Global scope of the CLEAR regional centres and global hub	387
Fig. 13.2	CLEAR theory of change phase II	389
Fig. 14.1	Theory of change underlying IDB group's work to support institutionalisation of evaluation in LAC	407
Fig. 14.2	Outreach of the office of evaluation and oversight	412
Fig. 16.1	Institutionalisation and evaluation use in the Americas	469
Fig. 16.2	Members in American VOPEs	487
Fig. 16.3	Correlation political and profession system	492

# LIST OF TABLES

Table 1.1	Dimensions of the institutionalisation of evaluation	19
Table 3.1	What role does civil society normally perform in evaluations?	77
Table 3.2	Do you think evaluation culture does exist in the country?	78
Table 3.3	Are evaluation reports (full version) available to the public?	79
Table 3.4	Are the results of current evaluations publicly discussed	80
Table 3.5	According to your experience and knowledge, do individual citizens, civil society organisations, private companies or other actors, require evaluations of, for example, political leaders?	82
Table 3.6	Are there university programmes of higher education for evaluators (Diplomas, Masters) in Bolivia?	83
Table 3.7	Are there other forms of academic or non-academic training on evaluation?	84
Table 3.8	Are there norms, guiding principles or others similar for evaluators in your country? Any guidelines developed by REDMEBOL that you know?	85
Table 3.9	Would you say that the evaluation market in your country is mostly dominated by independent, consulting or scientific research institutes?	86
Table 4.1	Normative basis on Monitoring and Evaluation in Brazil	100
Table 4.2	M&E Courses in Brazil	133
Table 5.1	Federal government commitment to evaluation	146

Table 6.1	Social programmes evaluated ex ante, MDSF Chile (2012–2018)	175
Table 6.2	Number of institutions evaluated and coverage. Period 2002–2017 DIPRES	180
Table 6.3	Budget evaluated. Coverage (1999–2017)	181
Table 6.4	Evaluation training programmes	192
Table 7.1	Evaluations carried out per year	206
Table 7.2	Instances evaluating service quality in Colombia	218
Table 7.3	Distribution by discipline	220
Table 7.4	Universities with indexed and specialised journals in evaluation	221
Table 10.1	Mexican evaluation legal framework	298
Table 10.2	National sectors with a greater improvement in evaluation	300
Table 10.3	Postgraduate programmes academic offer	313
Table 10.4	Diplomas and courses offer	314
Table 11.1	Ministry of Economic and Financial affairs. Number of evaluations of budget design and execution (EDEP) by sector/per year	331
Table 11.2	Ministry of Economic and Financial affairs. Number of impact evaluations (EI) by sector and status	332
Table 11.3	Ministry of Social Inclusion and Development. Number of evaluations carried out by type and per year	332
Table 11.4	Ministry of women and vulnerable populations. Number of evaluations carried out by type and per year	333
Table 11.5	Monitoring and/or evaluation postgraduate programmes offered to evaluators in Peru between 2016 and 2018	341
Table 11.6	Monitoring and evaluation specialisation courses offered to Peruvian evaluators between 2016 and 2018	342
Table 11.7	Training programmes and courses containing modules on monitoring and/or evaluation offered to Peruvian evaluators between 2016 and 2018	344
Table 13.1	CLEAR strategy 2013–2018	393
Table 13.2	Strategies for ECD	398
Table 16.1	Legislative institutionalisation of evaluation and evaluation use	454
Table 16.2	Sectors where evaluations are conducted	463
Table 16.3	Degree of spread across sectors where evaluations are conducted on a regular basis	464
Table 16.4	Institutionalisation of evaluation in the social system	482
Table 16.5	Professionalisation index	490

PART I

---

Introduction



# The Institutionalisation of Evaluation: Theoretical Background, Analytical Concept and Methods

*Wolfgang Meyer, Reinhard Stockmann,  
and Laszlo Szentmarjay*

## INTRODUCTION

The use of evaluation is increasing in a global scale. At least since the beginning of the twenty-first century, more and more countries are using evaluation as a tool for producing reliable scientific data grounding evidence-based policymaking. Despite all setbacks and populist ‘fake

---

W. Meyer (✉) · R. Stockmann · L. Szentmarjay  
Center for Evaluation (CEval), Saarbrücken, Germany  
e-mail: [w.meyer@ceval.de](mailto:w.meyer@ceval.de); [w.meyer@mx.uni-saarland.de](mailto:w.meyer@mx.uni-saarland.de)

Department of Sociology, Saarland University, Saarbrücken, Germany

R. Stockmann  
e-mail: [r.stockmann@ceval.de](mailto:r.stockmann@ceval.de)

L. Szentmarjay  
e-mail: [l.szentmarjay@ceval.de](mailto:l.szentmarjay@ceval.de)

© The Author(s), under exclusive license to Springer Nature  
Switzerland AG 2022

R. Stockmann et al. (eds.), *The Institutionalisation of Evaluation in the Americas*,

[https://doi.org/10.1007/978-3-030-81139-6\\_1](https://doi.org/10.1007/978-3-030-81139-6_1)



news', this is the main global trend and evaluation has become an important element of collective action in a broad variety of different areas such as development, education, environment or public health policy.

Evaluation is merely used for three purposes: in programme management, evaluation is used to learn for programme planning, the implementation of measures and the outcomes achieved by them. Policy evaluations contribute to good governance by delivering information on policy impacts and its sustainability and help to improve the legitimacy and credibility of politics. A public discourse about evaluation results may enhance the quality of dialogues on sociopolitical developments and support enlightenment of society (Stockmann, 2013, p. 74).

The increasing practical use of evaluation motivated more and more scientists for research on evaluation. This finally led to the present research project and its intention to generate an 'Evaluation Globe', covering countries throughout the planet and systematically compare the institutionalisation of evaluation on national level. The first book on **Europe** was published in 2020 (Stockmann et al., 2020) and this second book is on the **Americas**, both North- and South America. Two further volumes are planned on **Asia-Pacific** and **Africa** within the next years.

The target is a descriptive one: institutionalisation of evaluation in three subsystems (political system, social system and system of professions) will be presented in a comparative way. In this volume, 35 authors report about eleven countries in the **Americas** and three transnational organisations important for this region.

While the first volume on **Europe** already presented the general purpose and theoretical background, this chapter will focus on the **American** specifics and some further methodological details. Nevertheless, the next part of the chapter will give a brief overview on the theoretical background of the Globe-project. It will also include some remarks on comparative research and the principles followed here.

The following part offers a description of the main indicators and the research questions that will be answered by the authors. The concept was developed by the editors and discussed at the European Evaluation Society (EES) conference in 2016 with international experts and some of the authors of the **European** volume. Minor adaptations and additions for the Americas volume had been discussed at the IDEAS-RELAC-REDLACME conference in Guanajuato, Mexico, in summer 2017. A particular focus will be set on these changes and its consequences for national comparisons. Further remarks will focus on the selection process of countries and authors as well as on the quality assurance process managed by the editors.

The presented case studies and the contributions about transnational organisations are framed by three contributions from the editors: (1) an introductory chapter on the theoretical framing of the research project, (2) a synthesis assessing the findings from the Americas and (3) a comparison between the findings from Europe and the Americas.

## STATE OF RESEARCH AND THEORETICAL BACKGROUND

### *State of Research on the Institutionalisation of Evaluation*

As already shown in the **European** volume, the first comparative overview on the institutionalisation of evaluation in 21 countries and three international organisations was published by Furubo et al. (2002) and updated by Jacob, Speer and Furubo in 2015. The authors used nine indicators which had been rated by country experts. Almost all countries were from **Europe** and **North America**, none of the **Latin American** countries was included. For the development of evaluation, the authors identified three internal drivers—political constellation, financial situation and constitutional specifics—and one external driver—pressure from donor countries or organisations (Furubo et al., 2002; Jacob et al., 2015).

Two studies with a more global perspective had been mentioned in the **European** volume: in 2013, Barbara Rosenstein did an Internet research in 115 countries to analyse national evaluation policies. She found 20 countries with such kind of evaluation policies and 34 countries without such policies but with evaluation routines (Rosenstein, 2013). Eight of this 54 countries were from the **Americas**: **Argentina, Brazil, Canada, Chile, Colombia, Costa Rica, Mexico** and the **USA**—with the exception of Argentina and Brazil they had well-developed national evaluation policies, more countries than in Europe. All these countries are also included in this volume and one can find further, more actual and extended information on policies and regulations on evaluation in this book.

In 2016, Reinhard Stockmann and Wolfgang Meyer edited a reader on the ‘Future of Evaluation’ where more than 30 authors from 20 countries offered an overview on the professionalisation of evaluation in their countries (Stockmann & Meyer, 2016). This book is on differences in development paths, especially whether there is a tendency for global convergence towards one commonly shared evaluation culture or whether there is an increasing differentiation of evaluation following national or sectoral political cultures. Although this reader offers a lot of interesting

insights in many different countries, it is not a systematic analysis of the institutionalisation of evaluation.

Beside these studies, there are also some studies with a particular focus on the **Americas**. In 2005, the World Bank in partnership with the Inter-American Development Bank organised a regional conference on monitoring and evaluating the performance of public programmes. Political representatives from eleven countries participated at this conference which shared experiences from five **Latin American** countries (**Brazil, Chile, Colombia, Mexico and Peru**) and included results from a comparative study in 29 OECD (Organisation for Economic Co-operation and Development) member states (OECD, 1997). As one of the main results mentioned in the proceedings, “the approaches to linking with decision-making processes differ considerably, reflecting their different approaches to public sector management and resource allocation” (May et al., 2006, p. 60). Neither the quality of the evaluation nor the existence of an evaluation policy is supporting the use of evaluation for decision-making. In general, evaluation has become a political issue and “thus institutionalisation of the monitoring and evaluation framework is essentially a question of how to set and enforce a structure of incentives and disincentives” (May et al., 2006, p. 70).

The CLEAR-Centre (Centers for Learning on Evaluation and Results) in Mexico published a reader on **Latin America** in English this year (Pérez-Yarahuán & Maldonado, 2020), however, it presents the results which had been published in 2015 already in Spanish language. It contains studies from ten **Latin American** countries (**Argentina, Brazil, Chile, Colombia, Costa Rica, Ecuador, Mexico, Peru, Uruguay and Venezuela**) and states “a wave of dynamism in the creation of institutions associated with the issues of monitoring and evaluation of public programmes and policy” in **Latin America** (Pérez-Yarahuán & Maldonado 2020a, p. 15). For a systematic and comparative analysis, the editors stated four key dimensions to be investigated by the country authors: (a) formal recognition of monitoring and evaluation by the state, (b) existence of planning for these monitoring and evaluation activities, (c) explicit evaluation methodologies regulated by rules and formal procedures and (d) instrumental, symbolic and conceptual use of evaluation and its information gathered (see for further details: Pérez-Yarahuán & Maldonado 2020a, pp. 20f.). The findings support the assumption “that systems must interact in a dynamic way with the political and sectoral environment, and their components may follow different tracks determined

by their national context” and “that there will be no absolute convergence or isomorphism in the institutional configuration of these public management functions” (Pérez-Yarahuán & Maldonado 2020b, p. 390).

The CLEAR-study comes close to the approach followed in the Globe-project and the results offered support several of the findings presented in this reader. However, there are some differences in the Globe-project concept that will help to provide further insights. Although the English version of the CLEAR-study is just recently published, the results are about ten years old and there is a fast development of evaluation in **Latin America**. Moreover, the CLEAR-study does not include **North America** and the North–South comparison is a clear addition of the Globe-project.

Two other differences should be highlighted here to emphasise the additions of the Globe-project: first, the CLEAR-study is limited to evaluation as part of the national political system. The Globe-approach does not deny the importance of the political system for the development of monitoring and evaluation systems on the national level but goes beyond this as will be shown later. However, as stated for the key result of the investigations on Europe “above all, the development of evaluation in Europe has been policy driven” (Stockmann et al., 2020, p. 518). The CLEAR-study offers a certain indication for a similar role in **Latin America**, but it does not deliver any material for the development of monitoring and evaluation in the social system.

There are three reasons why the social system should be included in research on institutionalisation of evaluation and why it is included in the Globe-project:

- (a) the civil society acts as an important ‘watchdog’ for the political system, including critical perceptions of traditional (‘the fourth estate’, e.g. Schultz, 1998) and digital (‘the fifth estate’, e.g. Flew & Wilson, 2012) mass media. In a democratic system, the critical assessment of government by its citizens is essential.
- (b) the civil society organises itself and provides public services more or less independent of the state (e.g. Rosenbaum, 2006). This may generate a different understanding of monitoring and evaluation and lead to different forms of institutionalisation.
- (c) the offer of capacity building services is not a state task, although it is often provided by state organisations (schools, training centres, universities). In general, this task is part of the system of professions

(e.g. Abbott, 1988) that provides production, exchange and education of specialised knowledge. Due to its importance for offering evaluation services, the Globe-project put a particular focus on this system, in difference to the CLEAR-study, and completes the analysis on the evaluation market.

Second, the CLEAR-study does not follow a theoretical approach as the editors stated (Pérez-Yarahuán & Maldonado, 2020a, p. 16). This is the most important difference to the Globe-project that is based on a clear theoretical concept, described in the next section.

### *Theoretical Background of the Globe-Project*

From its early beginning in Sociology, institutions are seen as important bridges between individuals and society (Stachura et al., 2009; Traugott, 2013). Especially the structural functionalism in the 1960s highlighted the stabilising function of institutions and used this way of thinking among others for a sociology of professions (Brante, 1988; Parsons, 1939). As Talcott Parson stated it: “[the] study of the institutional framework within which professional activities are carried on should help considerably to understand the nature and functions of some of these social ‘constants’” (Parsons, 1939, p. 457).

While structural functionalism was criticised for overestimating the stability of social systems, modernisation theory emerged as the most important approach for explaining social change. Driving force is the horizontal process of social differentiation, that installs specialised functional subsystems (like politics, law, education, health, etc.) and links them interactively by institutions in a division of labour (Luhmann, 1983, 1988). The need for differentiation and specialisation is a result of intrinsic developments like technological progress, changing framework conditions like for instance climate change or deriving new individual desires like changing consumption patterns. These processes put the existing subsystems under pressure and they try to stabilise themselves by adaptation of their institutions (e.g. Alexander, 2001).

According to the theory of social differentiation (Schimank, 1996), social change is a continuous modernisation process that increases the complexity of social (sub-)systems and the adaptability of social institutions. Several different subsystems have been analysed by modernisation theory, with a particular focus on the development of its institutions (for

instance Giddens, 1996; Inglehart, 1998; Lerner, 1968; Parsons, 1971; Zapf, 1991). In general, modernisation theory emphasises the interrelations between key institutions especially in the economical (markets) and political (democracy) subsystems, interwoven with national organised social systems and its individual ownership and participation rights (Pollack, 2016).

Institutions can be defined as a set of rules and norms to regulate social behaviour within a given social (sub)system and they are interlinking different levels of society from the individual micro to the global macro-level (Peters, 2019, p. 158). Therefore, institutions are not only in the focus of macro-analysis like modernisation theory but also on the meso-level as for example in organisation theory (e.g. Coase, 1937) or at the micro-level in rational choice theory. In an overarching perspective, March and Olsen (1984) highlighted the relative autonomy of political institutions not only as an orientation for collective action but also for sensemaking and sharing a common concept of governance. It is also about the options for shaping social development and steering it towards the favoured direction. Rules and governance systems are necessary to define system borders and to regulate cross-border traffics (Kapitanova, 2013, p. 257). Such kind of institutions include formal and informal control systems for opening and closing system borders, for regulating exchange between subsystems to guarantee system integration and to supervise transactions (Armingeon, 2016). This is also the entry point for research on institutions in political sciences, offering a broad variety of approaches and a long history (Peters, 2019). Interdependences between subsystems are primarily power relations and therefore key for understanding modern societies, their conflicts and governance processes (Schimank, 2001).

For the Globe-project, the following objects are identified to be key for analysing the institutionalisation of evaluation (Meyer et al., 2020, p. 12):

- “Rules, norms and regulation on evaluation, implemented in the already existing social subsystems. [...]”
- Evaluation processes, procedures and routines, implemented within a broad set of organisations or networks at least as a possible way of practice within a certain policy field. [...]”
- Finally, institutionalisation is a process with certain steps and it is probably a long way toward building a complete and deep-seated institutional framework. [...]”

### *Different Institutional Frameworks in the Americas*

The institution theory chosen as the basis for our evaluation approach is a general theory and should be valid in all kinds of societies. However, as is true for all theories, the respective contextual conditions play a central role. Therefore, it makes sense to show the differences in the institutional framework between **North** and **South America** and between the **Americas** and **Europe**. First of all, it is important to note that the colonisation of the **Americas** is merely a result of policies of expansion in **European** colonial states and this led to a transfer of political cultures and institutions from the ‘home countries’ to new build political units in the **Americas**. As Mahoney (2010, p. 1) stated in his comparative history on **Latin America**, “the institutions established during colonialism [...] exhibited over-time effects, whether directly through their own persistence or indirectly through the actors and processes that they brought into being”.

That is, **South** and **Central America** were influenced by Spain and Portugal and its feudalistic institutions, while **North America** was mainly colonised by the British and the French and was later characterised by several waves of immigration from a large number of European states who fled from poverty, war, religious and political persecutions (McKeown, 2004, p. 156). Unlike the colonisation of **North America** the Spanish institutions supported the integration of indigenous people as ‘Spanish citizens’, but this does not protect them from exploitation and cultural subjugation. In addition, the Spanish and Portuguese culture, language, religion and not least the feudalistic system of the colonisers have shaped the institutional framework in **Latin America**. Whereas in **Latin America** there are a number of states in which you find a high proportion of Indians and political and social structures are determined or at least influenced by them, the colonisation of **North America** led to the almost complete eradication of the indigenous Indian population. In difference to the development of the **South**, the nation-building in the **North** was more than a ‘melting-pot’, forming strong and huge independent nation-states out of much more heterogenous migration societies with a dominant role of democratic political and stabilising institutions. As a result, this led to economic prospering and politically more open societies in the **North**, while most **Latin American** countries suffered by economic dependency from **Europe**, inequality caused by the feudal system and instable political developments (Bulmer-Thomas, 2003).

If one compares the institutional framework in **Europe** with that of the **Americas**, a much greater diversity of structures that have developed over many centuries can be seen in **Europe**. The countries of the **Americas** naturally lack comparable long-term development processes and are not so strongly shaped by individual national cultures. In comparison with **Europe**, with regard to the institutional framework, it can also be observed that **North America** has succeeded in establishing two large nation states (**USA** and **Canada**), which on the whole have a greater homogeneity than **Europe**.

**Latin America** is characterised by a variety of independent nation states, without, however, having established a federation of states comparable to the European Union, as is the case in **Europe**.

Trials like the OAS (Organisation of American States) or the CELAC (*Comunidad de Estados Latinoamericanos y Caribeños*) are far behind the political institutionalisation processes at the **European** level. In general, both transnational entities lack of a commonly shared vision and still stuck in national egoisms and ideological differences (Segovia, 2013, p. 105). Moreover, they do not create common action and joint households for regional development or any other form of activities for coherency policies.

There are certain activities of development cooperation, including bilateral as well as multilateral action programmes or transnational organisations with branches in **Latin America** and the **Caribbean** (and sometimes their headquarters in North America). There is certainly no transnational institution with a comparable dominant role like the EU in **Europe**. The selected transnational institutions—**Interamerican Development Bank**, **Independent Evaluation Group** of the World Bank and the **CLEAR Centre** in Mexico—are important actors in the field of evaluation both in financing development projects as well as initiatives for improving evaluation in the region (e.g. by capacity building activities). In general, this is a ‘North–South’ transfer in the **Americas** and it is one of the key questions to be answered here whether these transnational institutions are successful in supporting an evaluation culture especially in **Southern America**.

What do these framework conditions mean for the institutionalisation of evaluation? There are three effects that might be considered in this volume:



- (a) The missing of a dominant transnational institution like the EU that proofed to be very important for the spread of evaluation in ‘latecomer states’ like **Central Eastern Europe, Spain** and **Portugal**. Cohesion policy is one of the most important drivers in Europe due to its huge financial impact and its linkage to obligatory evaluations. It is an interesting research question whether transnational development organisations are able to play a comparable role in the **Americas**.
- (b) The strong division between **North** and **South America** leads to the question whether one can still find more differences between these regions than between the countries. Is there a particular ‘Latin American-culture’ of evaluation with different forms of institutions that can be separated from the North American- (and European-) development? This will be explored in a comparative chapter including also the results from the Europe volume (see Chapter 17 in this book).
- (c) By taking three different systems—the political, the social and the professions system—into the focus, there might be varying effects of the institutional history briefly described here. A comparative analysis of the three systems should also look for similarities and differences in **North** and **South America** as well as in **Europe**.

## GLOBE-APPROACH, METHODOLOGICAL CHALLENGES AND DECISIONS

As briefly described above, monitoring and evaluation can be seen as an essential part of social institutions for steering societies as a whole on global, national or local level and within subsystems like education or economy for various forms of corporate governance in profit and non-profit organisations. The research on this is the task of the political economy of governance (Schofield & Caballero, 2015; Payne & Phillips, 2014; Bertelli, 2012; Barker, 2010; Ebener, 2008). The focus of the Globe-project is laid on the national level and its governance institutions, so the first object of analysis is the political subsystem and its governance tasks.

### *Institutionalisation of Evaluation in the Political System*

The basic principle of modern democratic political system is the separation of powers and the interwoven interdependences between executive, legislative and judiciary. The formal legal framework includes laws, acts, decrees, edicts and all kind of regulations. Constitutional organs defined here (e.g. the administration, the police, lawyers, etc.) have to follow this legal framework strictly without assessing or interpreting on behalf of their own interests. Responsible for the formulation and definition of the legal framework is the legislative, in most countries the elected parliament. For enacting, a majority of votes in parliament is needed and parties (or other forms of coalitions or associations) develop ideas and concepts communicated publicly and in negotiations with other parties. From a methodological perspective, measuring this legal framework is easy to analyse because all kind of regulations have to be published and even the enacting process is well documented. Nowadays, these documents are digitalised and can be accessed by everyone. There are only two difficulties left: the broad variety of regulations with shared responsibilities of different administrations (Lowe & Potter, 2018) and the use of the term evaluation in regulations. The instructions of the Globe-project include a glossary of key terms to guarantee a shared understanding and treatment of all authors. The research is limited to the national level and universal rules, more specialised and narrowed norms should only be mentioned if they have a broader importance.

Formal legal frameworks are important preconditions for collective administrative action—the executive needs guidance from the legislative for compliant activities. But the legislative is also dependent of the executive: without implementation all kind of rule-setting is senseless. The administration has to accept the rules as legitim orders and use it for routines of collective action (Goodnow, 2003). Evaluation rules are necessary but not sufficient for societal relevance—to contribute to social betterment, evaluation has to be institutionalised into regular administrative processes (Kettl, 2018). Things are getting more complex if other actors outside the political system like private companies or civil society organisations should be involved into governance to increase effectiveness of regulations (Mayntz, 2003). Such kind of ‘public–private partnerships’ need new forms of institutions for rational governance as well as a customer-oriented public administration developed in concepts of ‘New Public Management’ (e.g. Christensen & Lægrid, 2007; Kettl,

2015; Ansell & Torfing, 2016). The shared idea of these reforms is to transfer ‘service- and quality orientation’ from private to public sector for improving effectivity and acceptance of bureaucratic governance (Stockmann, 2008, p. 57ff.). The methodological challenge for the Globe-project is the recognition of all relevant governance processes and the assessment of the role of evaluation within these processes. If evaluation is institutionalised regularly in all kind of governance systems in various policy fields, this should be documented and mentioned in the country analysis. This implies two different assessments: first, the governance process must be analysed and if evaluation is implemented here. This can be easy if there is a national evaluation policy with certain implementation rules, but—as in **Europe**—it may be difficult if there are sector-specific traditions and non-formalised routines. This leads to the second assessment on the scope of evaluation practice and the question in which policy fields evaluation is implemented. There might be huge differences about the formal state, the inclusion of different state and non-state actors, the role of control institutions like for instance the parliament, audit offices or other authorities, and the forms and instruments used for evaluation. In general, this must be an open question answered by in-depth analysis and good descriptions.

To summarise: there are four key elements of the political system in the focus of the Globe-project (see also Table 1.1, Column 1):

- National Acts, laws, decrees, rules and all kind of regulations that include evaluation as a task and national evaluation policies or routines that are regularly implemented in governance processes (row 1 ‘National laws, regulations and policies’),
- Embeddedness of evaluation in parliamentary or other (e.g. audit offices) control structures (row 2 ‘Parliamentarian and national audit structures’),
- Evaluation units as core elements of political and administrative organisations (row 3 ‘Organisational structures’),
- The use of evaluation results for decision-making or other purposes in the political process (row 4 ‘Evaluation practice’).

### *Institutionalisation of Evaluation in the Social System*

The use of evaluation results is not only limited to a direct instrumental use for political decision-making. Evaluation can, for example, also be used for control reasons and with the same intention like other controlling or accountability instruments (e.g. in performance-accountability regimes, Han, 2019). The use of evaluation results may also be limited for public administration purposes (and therefore to the political system), but it can also be used in a much broader sense. This is especially true for the North American political culture: as already described by Alexis de Tocqueville in his famous work ‘*De la démocratie en Amérique*’ (Tocqueville, 2003), participation of citizens within civil society is a key element of the democratic model derived in **North America** (see for the actual situation Edwards, 2020, pp. 17ff.). These are important roots for the political system of an open society (Cohen & Arato, 1994). Civil society is an institution for transferring interests from social to political system and for allocating responsibilities from the political to the social system.

There are two main functions of civil society organisations: the first one is a ‘watchdog function’ to control state activities being in the interest and for the need of the citizens. Evaluation results may be used for this task, either if they are commissioned by state authorities or by civil society organisations on their own behalf. In short: citizens can hold their governments responsible by using evaluation organised by civil society.

The second one is the service function to take over responsibilities for providing social services in addition or on behalf of state authorities (Howell & Pearce, 2001; Kalm & Uhlin, 2015; Laville et al., 2015; Lovan et al., 2003; Ojo & Mellouli, 2018). If the state authorises private non-profit organisations to overtake state duties, there is some need to hold these organisations accountable and evaluation may become an appropriate instrument for this duty (Carman, 2009; Cutt & Murray, 2000).

From a methodological point of view, the analysis of evaluation in civil society is difficult because it is not a closed and highly formalised system like the political system. It is very heterogenous and diverse instead with a broad variety of different and sometimes contradictory institutions, sometimes changing quickly according to social change or internal dynamics. The Globe-project emphasises national and general institutions, so best practice solutions or forerunner civil society organisations should not be the main topic of the country reviews.

To sum it up: for the institutionalisation of evaluation in the social system, four aspects are asked to be investigated (see also Table 1.1, column 2):

- Evaluation as an institutionalised or at least widely used instrument for improving civil society organisations (row 1 ‘Use of evaluation by civil society’),
- Evaluation results are commonly used in public discussions (row 2 ‘Public discourse’),
- Involvement of citizens, civil society organisations or other actors into public evaluations (row 3 ‘Participation of civil society’),
- Civil societies demand evaluations from political institutions (row 4 ‘Demand for evaluation’).

### *Institutionalisation of Evaluation in the System of Professions*

The final object of analysis is a very unique part of the social system that can be distinguished from civil society and its particular relationship to the political system. On one side, it is covering the ‘supply’ side of service markets because it offers education and training for achieving the skills and knowledge providing these specialist services. Universities—both state and private financed—provide study programmes and certificates to proof the ability for managing these tasks. In difference to occupations, study programmes are more closely linked to research (knowledge production) but less to practice (knowledge use). On the other side, the universities are independent both from the state and private profit or non-profit organisations. They are not developing study programmes primarily for the service market but for their own purposes in research. This constellation results in an own system of professions as described by Abbott (1988).

Knowledge transfer is one key element of a profession (Freidson, 2001), but it is not the only one. Research on profession offers several different perspectives and criteria to define professions and to distinguish them from non-professions (MacDonald, 1995). Some of these authors emphasise on social interaction as the main source for developing a ‘professional identity’ (Colbeck, 2008). Communication between members of the profession is necessary to shape its profile and to transfer experience

as well as practicable professional knowledge. Such kind of communication is installed within the academic system and a new profession has to be implemented into this existing academic communication system (Chan & Fisher, 2008; Schiele et al., 2012). The most important institutions for academic exchange are peer-reviewed journals (Larivière et al., 2015). However, communication within a profession cannot be limited to academia, most professions develop additional forms of communication especially organised by professional associations to bridge the gap between academia and practice (Friedman & Phillips, 2004; Greenwood et al., 2017).

The last element highlighted in professionalisation research is the degree of independence achieved by professions, seen in the ability to set their own rules and norms for professional practice—and to push them through at the market (Rüchemeyer, 1983). In general, this is a process of social closure, for instance by limiting the right to deliver services to people with a particular license to practice provided by the profession. The target of the profession is the monopolisation at the supply side of the market and the protection against other competing professions (MacDonald, 1985; Richardson, 1997). However, if one looks at existing standard systems or ethic codes developed by professions, they aim more on the practice within the profession and try to improve the quality of their services (Ingvarson, 1998).

From a methodological point of view, the aspects mentioned here are easy measurable because study programmes and academic journals, professional networks and their communication platforms, and standard systems or ethical codes are published and promoted to the broader public within a country. Nevertheless, some difficulties may occur according to the definition and understanding of evaluation as a subject of professions—especially if evaluation is implemented within existing professions and not as a profession by its own.

To summarise the task of the analysis of the system of profession, the following aspects should be in the focus of empirical research (see also Table 1.1, column 3):

- Existence of academic study, trainings and qualification programmes for evaluation (row 1 ‘Academic education and training practices’),
- Existence of academic journals or other media and fora for evaluation (row 2 ‘Journals and communication platforms’),

- Existence of professional organisations, networks or associations that manage the exchange about evaluation and promote the development of this discipline (row 3 ‘Professional organisations’),
- Existence of generally binding standards or rules for evaluation (row 4 ‘Existence of and compliance to standards’).

The Globe-project captures the institutionalisation of evaluation as different forms of implementation in three important systems. Evaluation can be used as an instrument of governance in the political system and the focus is set on general use, not limited to certain policy fields. National acts, laws, decrees and other regulations may assign how evaluation should be applied in administrative processes. Evaluation practice develops within this legal framework and several different forms of approaches and methods may be used in different policy fields. Evaluation results are supposed to be used for governance decisions and to improve the quality of state activities and public services.

While public services are not only delivered by state authorities but also by civil society organisations, the social system comes into the focus. Beside claims for evaluation by the state, civil society organisations may use evaluation on their own behalf to improve their own activities. This may, on one side, be targeting on the delivery of (public) services or, on the other side, be part of activities for controlling the state, aiming on public interests and improving state performance for these interests. Civil society and mass media may demand for evaluations commissioned by the state or to be involved in such kind of evaluations to enrich the results. They may also commission evaluations on their own behalf or use evaluation results for requests for public action.

The system of professions may include academic study programmes on evaluation at universities and provide opportunities to publish evaluation results and results on evaluation research in scientific journals with a corresponding focus. Additionally, there might be a communication platform for evaluation practice that is more useful for applications than for pure research. In many cases, such kind of exchange is organised by Voluntary Organisations of Professional Evaluation (VOPE), networks and association with members who are interested in discussions on evaluation. Finally, these organisations (or other actors) may provide quality standards or ethic codes on how evaluations should be conducted.

The beforementioned aspects are summarised in the following Table 1.1:

**Table 1.1** Dimensions of the institutionalisation of evaluation (own development)

<i>Institutionalisation of Evaluation in different Subsystems</i>		
<i>Political System: Institutional Structures and Processes</i>	<i>Social System: Dissemination and Acceptance of Evaluation in Society</i>	<i>System of Professions: Evaluation as a Discipline</i>
National laws, regulations and policies	Use of evaluations by civil society	Academic education and training practices
Parliamentarian and national audit structures	Public discourse	Journals and communication platforms
Organisational structure	Participation of civil society	Professional organisations
Evaluation practice	Demand for evaluations	Existence of and compliance to standards

### *Comparison, Selection of Countries and Methodological Procedure*

In general, the Globe-approach is a comparative one and the challenge is to find a shared scale for assessing the institutionalisation of evaluation in a comparative way. The challenge is composed both by a theoretical and a methodological component. There are certainly a lot of historical and cultural specifics that makes each of the countries and their institutional setting unique. For a foreigner, it is in most cases not easy to understand these country specifics and therefore it is a common way to team up with national experts for cross-country comparisons.

Cross-country comparative research implies some specific difficulties that must be addressed in the research process (for an overview, see Esser & Vliegenhart, 2017; Goerres et al., 2019; Hantrais, 2008; Keman & Pennings, 2017; Wong, 2014). The three most important challenges will be briefly mentioned here: construct equivalence, synchronising research methods and sampling procedures.

As the work on the **European** volume revealed, the concept of ‘evaluation’ is used in different forms and meanings across sectors and countries. Sometimes, the actors in a country—especially if they are not familiar with international evaluation norms and standards—mix up or link evaluation to other approaches like controlling, monitoring or quality management. Although there is a globally shared and communicated understanding



of evaluation, construct equivalence is not guaranteed for evaluation practice.

While the roots of evaluation are clearly in **North America**, the understanding of the concept is assumably more homogenous than in **Europe** or **South America**. Briefly said, the understanding of the term in **Latin America** is much closer to monitoring than in **North America** or **Europe** and this must be mentioned by comparisons. In general, this is the task of the country authors and therefore construct equivalence is not only linked to national definitions or practices but also to the researchers itself. While there is a clear and commonly shared understanding of evaluation at a global scale (due to the high degree of exchange and organisations such as the International Organisation for Cooperation in Evaluation—IOCE), this does not cause difficulties—neither in **Europe** nor in the **Americas**. Furthermore, the methodological solution for the Globe-project is a shared glossary and no queries occurred about this. This had been proofed by a quality assurance process described below.

Synchronising research methods is, on one hand side, easy because some of the research objects are well-documented, publicly assessable and easy to trace. To give an example: while the IOCE is a global umbrella organisation of all VOPEs, the existence of such kind of organisation, network, association et cetera in a particular country is easy to trace by the published list of this organisation. The registration on this list is free and IOCE offers some organisational help for newcomers if needed. All large VOPEs are members of the IOCE and it is rather unlikely that an existing VOPE is not on this list. But even if this is the case it should be easy to find this VOPE if one is doing a research on the national evaluation community. Such kind of information are not depending on elaborated research methods and easy to proof.

However, on the other side, there are some objects that are very difficult to analyse. The assessment, for example, of evaluation practice is depending on the publication practice that may also differ between policy fields. Strategies to explore this and to produce appropriate evidence on the usual practice cannot be standardised due to this broad variety of existing forms of treatment. Therefore, it is not possible to preset a particular methodology and to monitor a similar procedure across different countries. The only opportunity is to accompany the chosen scientists very closely and to discuss methodological issues quite frankly. In some cases, this may balance against comparability—but the descriptive task is

of higher importance than the comparison. This must be considered at certain parts of comparative interpretations.

Finally, the selection of scientists is important and one of the sampling tasks within the Globe-project. Another one—and closely linked—is the selection of countries for this volume. Selecting the right cases—in this case the countries and transnational organisations as well as authors—is always a challenging task.

The main aspect for the selection in this context were contacts and networks which already existed before between the editors and evaluation and research experts within the **Americas**-region. This is especially the case for joint research and evaluation projects in the past, the since 2008 existing universities cooperation between the Saarland University (UdS) and the University of Costa Rica (UCR), which had been extended 2017 to the Pontifical Catholic University of Ecuador (PUCE), or activities in the context of the elaboration of the ‘Evaluation Standards for Latin America and the Caribbean’ of ReLAC.<sup>1</sup>

A second aspect was through the presentation of the overall concept and the interactions on thematic conferences, like the European Evaluation Society conference in 2016 or more important the IDEAS-ReLAC-REDLACME conference 2017 in Guanajuato/Mexico, where a special talk was held on the topic ‘*Evaluation Globe – Compendio sobre la Institucionalidad de la Evaluación*’ and the concept was presented and discussed the very first time to and with the **Latin American** evaluation community.

The third aspect was literature and Internet-based research as well as further talks with experts to gain insights into country-specific contexts where no connections from editors’ side existed before.

The aim of these three aspects was, first, to verify from an **American** perspective where evaluation does play a significant role within the countries and where it does not, against the elaborated analytical framework (see Appendix 1) of the editors, and second, to find appropriate authors.

At the end, eleven countries (**Argentina, Bolivia, Brazil, Canada, Chile, Colombia, Costa Rica, Ecuador, Mexico, Peru, and the USA**) as well as three transnational institutions (**CLEAR-LAC, Independent Evaluation Group (IEG)** of the World Bank Group and the **Office of Evaluation and Oversight (OVE)** of the Inter-American Development Bank) had been selected to be included in the present volume on the

<sup>1</sup> <https://www.relac.net/biblioteca/evaluation-standards/>.

Americas (Fig. 1.1). This corresponds to a pool of 35 authors who took over the herculean task and present their country- or organisation-specific landscape of the institutionalisation of evaluation.<sup>2</sup>

Of course, it was not possible to include all countries of the **Americas**. For example, it would have been important to have at least one country from the Caribbean-region involved. However, unfortunately, it was not possible to find someone, who was willing to take on this task.

Just as in the **Europe** volume, all the authors for the **Americas** volume had been equipped with—and, of course, asked to commit themselves to it—the necessary research materials to conduct their country-specific case studies, which included the general project concept with the theoretical background presented above, an in-depth analytical framework, a glossary, to secure a common understanding of used key terms and concepts, and a manual of production, to meet the production criteria of the publishing house.

The analytical framework had been developed to be a lean and feasible instrument, on the one hand to keep the efforts for the experts within an acceptable frame, on the other hand to guarantee that the case studies are focused on the most relevant aspects and still keep an acceptable length.

In comparison with the **Europe** volume, only one minor adaptation in the analytical framework had been made, to maintain the level of comparison of the—to be published—four volumes: with regard to the VOPEs it was necessary to further distinguish between different levels of organisation. This means, that special focuses should be given first, to the degree of organisation—which means to differentiate between open networks that anyone can join, or closed associations that can act as formalised entities vis-à-vis other actors—and second, their organisational power—which means the actual number of members.

With view on their proceedings, the authors were first asked to realise co-authorships by themselves, which had several reasons: not everyone is a specialist in every of the three subsystems to be presented in the case studies. Therefore, further country experts from the specific field should be called in. This gave them the opportunity to split work packages, accelerate the production process, as well as—and in such kind of international

<sup>2</sup> At the end of 2019, it was necessary to change the author teams for **Brazil** and **Colombia**. A special thanks goes therefore again to the here present authors of both countries, to realise their case studies in a quite shorter time frame than the other authors had.



**Fig. 1.1** Selected countries and transnational organisations for the case studies (Own development)

project very relevant—to promote cooperation, which was taken on by many authors in a very positive way and created—at least from the editors' perspectives—very fruitful and collaborative working atmospheres.

Second, to create content for their case studies they were asked to do extensive literature and document analysis, to answer all relevant points which were raised by the analytical framework. Finally, they were given free hand for the conduction of interviews or surveys with different evaluation experts from within their countries or organisations, for example to fill up gaps which could not be answered via the literature and document analysis.

The whole production cycle stretched from the search and selection of the cases as well as of appropriate authors and very first meetings with them in 2017 to the finalisation of the manuscript at the end of 2020. Central during this period was the continuous support of the authors by the editors in order to comply with the methodological, content-related and time-related requirements of the publication. This included especially the quality assurance of the submitted chapters (respectively their drafts), guaranteeing the following to the analytical framework and, with this, secure the comparability of all the country studies throughout the four volumes.

In the majority of cases, a two- or three-stage review process was conducted to realise this goal and capture (almost) all details of the analytical framework. This review process was conducted internally by the editor's institution Center for Evaluation (CEval) at the Saarland University/Germany and was done by highly experienced evaluation and scientific experts in a four/six-eyes principle.

A special condition was that the majority of case studies were submitted and internally reviewed in—self-evident, when one does conduct a project in Latin America—Spanish language. Fortunately, the human resources at CEval did match these conditions. Without these language—and of course also cultural—capabilities, the access to the majority of country experts would had been—maybe—closed and the here presented volume less informative.

Some last words shall be allowed regarding the dissemination of the results: of course, the here presented published volume is the main result of the whole project. Nonetheless, it is also only a part of it, as the results were and will be used for further research and teaching, as well as presentations at different conferences. This was done in 2020 for example in the

context of the virtual formats of the gLOCAL Evaluation Week<sup>3</sup> in June or in September in a Mini-Series of the globally well-recognised International Programme for Development Evaluation Training (IPDET).<sup>4</sup> Some further presentations were planned for the same year, like in the framework of the annual conferences of the American Evaluation Association (AEA) or the Canadian Evaluation Society (CES), but had been refused due to global COVID-19 pandemic. But, as a proverb, says: *‘Postponed is not abandoned’*. For 2021 further presentations are planned for the gLOCAL Evaluation Week, IPDET, potential ReLAC-, AEA- and CES-conferences, events in the context of the universities cooperation between UdS, the UCR and the PUCE, as well as journal articles, with the aim which connects the editors as well all contributing authors: to enable cross-national learning for an interdisciplinary audience to better understand the institutionalisation of evaluation in different countries, regions as well as sectors.

## APPENDIX I.I ANALYTICAL FRAMEWORK: COMPENDIUM ON THE INSTITUTIONALISATION OF EVALUATION

<sup>3</sup> <https://glocalevalweek.org/>.

<sup>4</sup> <https://ipdet.org/>.

## *I. Political System: Institutional Structures and Processes*

<i>I.1 Evaluation regulations</i>	<i>I.2 Evaluation practice</i>	<i>I.3 Use of evaluations</i>
<ul style="list-style-type: none"> <li>• Are there national laws or regulations about evaluation or use of evaluation? If yes, which?</li> <li>• Are there sectoral laws or regulations about evaluation or use of evaluation (e.g. a law about school evaluation or evaluation in the higher education system as example for laws in the educational sector)? If yes, which?</li> <li>• Are there policies or strategies about evaluation or use of evaluation, either national or sectoral? If yes, which?</li> </ul>	<ul style="list-style-type: none"> <li>• With regard to the whole country: How would you describe the scope of conducted evaluations? Is it possible to speak of a frequent rhythm of evaluations, for instance for every new legislation or for every national program? Or is it rather non-specific? Does evaluation take place in all sectors/policy fields of a country (instead of only in the field of development cooperation for example)? And within one sector, is evaluation applied for measures funded in different ways or maybe only the ones that received funding by the European Union?</li> <li>• With regard to the whole country: How would you describe the relation between internal and external evaluations? Which form is carried out more often and for what purposes?</li> <li>• What are possible reasons for this (e.g. determination in laws, policies or regulations)?</li> </ul>	<ul style="list-style-type: none"> <li>• Which sectors are ‘good performer’ regarding use of evaluation and evaluation findings? Please describe up to 3 sectors that can be considered as leading in the field of evaluation’s use.</li> <li>• Which sectors are ‘bad performer’ regarding use of evaluation and evaluation findings? Please describe up to 3 sectors that are lagging behind in the field of evaluation’s use.</li> </ul>

(continued)

(continued)

<i>I.1 Evaluation regulations</i>	<i>I.2 Evaluation practice</i>	<i>I.3 Use of evaluations</i>
<ul style="list-style-type: none"> <li>• Are there administrative regulations about evaluation or use of evaluation in different policy fields (instructions, guidelines, etc.)? If yes, which?</li> <li>• What is the content of these laws/ regulations/ policies/ strategies or administrative regulations regarding independence of evaluation, quality, impact orientation and available budget?             <ul style="list-style-type: none"> <li>– Is use of evaluation specified? If yes, how?</li> <li>– How binding are specifications regarding use of evaluation?</li> <li>– What are aspired functions of evaluation (e.g. Planning and Steering, Accountability and Legitimacy, Enlightenment)?</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Is this relation differing with regard to sector or state level?</li> <li>• With regard to the whole country: How would you describe the relation between process and impact/outcome evaluations? Which form is used more often and for what purposes?             <ul style="list-style-type: none"> <li>– What are possible reasons for this (e.g. determination in laws, policies or regulations)?</li> <li>– Is this relation differing with regard to sector or state level?</li> </ul> </li> <li>• Does an independent evaluation institute exist in your country?</li> </ul>	<ul style="list-style-type: none"> <li>• On which aspect do most evaluations focus in these sectors (e.g. Planning and Steering, Accountability and Legitimacy, Enlightenment)?             <ul style="list-style-type: none"> <li>– In the case of different findings in different sectors: What might be possible reasons for these differences?</li> </ul> </li> <li>• Which professional groups use evaluation and evaluation findings regularly (e.g. political decision makers, program or project manager, administrative staff)?             <ul style="list-style-type: none"> <li>– For what reasons are evaluations and/ or evaluation findings used by these groups?</li> <li>– How is the use of evaluation findings guaranteed (for instance: management response mechanisms, implementation of monitoring for evaluation results, others)?</li> </ul> </li> </ul>

(continued)



(continued)

<i>I.1 Evaluation regulations</i>	<i>I.2 Evaluation practice</i>	<i>I.3 Use of evaluations</i>
<ul style="list-style-type: none"> <li>• Is evaluation and use of evaluation findings embedded in parliamentary structures? If yes, how?               <ul style="list-style-type: none"> <li>– Do parliamentarians in your country deal with evaluation findings for their own political work? If yes, to what extent (how often/ how detailed do they use evaluation findings)?</li> <li>– Do parliamentarians in your country demand evaluations for their own political work? If yes, to what extent? (How often? Do they commission evaluations? Do they publicly demand evaluations)?</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>– With a national responsibility?</li> <li>– With a responsibility for a specific sector or policy field?</li> <li>– Do independent internal departments exist, in ministries or elsewhere?</li> </ul>	<ul style="list-style-type: none"> <li>– Are there differences with regard to different sectors?</li> <li>• How is the quality of evaluations guaranteed (e.g. regular conduction of meta-evaluations analyses, competence requirements for evaluators, quality requirements for evaluations)?</li> </ul>

## *II. Social System: Dissemination and Acceptance of Evaluation in Society*

---

### *II.1 Institutionalised use of evaluations by civil society*

- Is it usual practice in your country that evaluations are used to provide knowledge for referenda or political decision-making on a communal basis?
  - If yes, how regularly does this happen? If not, what might be possible hindering factors?

### *II.2 Public perception and discussion of evaluation and evaluation findings*

- How well known is the instrument of evaluation in society?
- Are evaluation reports (full version) made publicly available?
- Is the general use of evaluation publicly discussed in media (benefits of evaluation, quality of evaluations and professionalisation of evaluation)?

### *II.3 Civil societies demand evaluations*

- Do individual citizens, civil society organisations, private enterprises or other actors in your country demand evaluations, e.g. from political decision-makers?
  - If yes, how often does this happen and under which circumstances/for what reasons? If not, why not? What might be possible hindering factors?

(continued)

---

(continued)

<i>II.1 Institutionalised use of evaluations by civil society</i>	<i>II.2 Public perception and discussion of evaluation and evaluation findings</i>	<i>II.3 Civil societies demand evaluations</i>
<ul style="list-style-type: none"> <li>• Are evaluations and evaluation findings used by individual citizens/ civil society organisations and/or private enterprises or other actors?               <ul style="list-style-type: none"> <li>– If yes, for what reasons (e.g. enforcement of their interests, knowledge or proof for work related issues, knowledge or proof for voluntary activities, etc.) and how regularly? If not, what might be possible hindering factors?</li> </ul> </li> <li>• Is it usual practice in your country that citizens or civil society organisations (NGOs, CSOs, churches, etc.) are participating in evaluations (as stakeholder)?               <ul style="list-style-type: none"> <li>– If yes, how regularly does this happen? What are different forms of participation (e.g. as interview partners, as clients, as users of evaluation findings, etc.)? If not, what might be possible hindering factors?</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>– If yes, to what extent? If not, what might be possible hindering factors?</li> <li>• Are findings of actual evaluations publicly discussed (surprising findings, different possibilities of dealing with these findings)?               <ul style="list-style-type: none"> <li>– If yes, to what extent? If not, what might be possible hindering factors?</li> </ul> </li> </ul>	

### *III. System of Professions: Evaluation as a Discipline*

<i>III.1 Academic study courses, further training et cetera</i>	<i>III.2 Profession/discipline</i>	<i>III. 3 Compliance to standards and quality obligations</i>
<ul style="list-style-type: none"> <li>• Do programmes of higher university education for evaluators (Diploma, Master) exist in your country? If yes, how many and where?</li> <li>• In which other scientific disciplines is evaluation instructed as scientific subject? Please give as many examples as possible.</li> <li>• Do other forms of academic or non-academic training exist? (e.g. e-learning, training by consultancies, else)?</li> </ul>	<ul style="list-style-type: none"> <li>• Which professional journals, newsletters or other ways/ media of communication (e.g. e-Mail or discussion lists) exist?</li> <li>• Which professional journals from other scientific disciplines deal with evaluation regularly?</li> <li>• Does a professional organisation (VOPE - Volunteer Organisations for Professional Evaluation) exist in your country?               <ul style="list-style-type: none"> <li>– How is it organised (closed associations or open network)?</li> <li>– What is there actual number of members?</li> </ul> </li> <li>• Do standards, guiding principles for evaluators or s.th. similar exist in your country?               <ul style="list-style-type: none"> <li>– Developed by the VOPE?</li> <li>– Adopted from another VOPE?</li> </ul> </li> <li>• Would you say that the evaluation market in your country is mostly dominated by freelancer (people calling themselves evaluators), consulting firms or scientific research institutes?</li> <li>• Does a certification system for evaluators exist in your country?</li> </ul>	<ul style="list-style-type: none"> <li>• Do professional organisations ask their members to follow standards or guiding principles? If yes, how obligatory is this?</li> <li>• Do clients demand a certain evaluation quality and/ or compliance to standards? How does this demand look like (is it obligatory)?</li> <li>• To what extent do evaluators (and clients) follow these standards and/ or quality obligations?</li> </ul>

(continued)

(continued)

---

*III.1 Academic study courses, further training et cetera*

---

*III.2 Profession/discipline*

*III. 3 Compliance to standards and quality obligations*

---

- Does an authority, which might be asked to conciliate in case of a dispute, arbitration board exist in your country, like an arbitration board or ombudsman?
  - Does a professorship for evaluation exist in your country?
-

## REFERENCES

- Abbott, A. (1988). *The system of professions. An essay on the division of expert labour*. Chicago University Press.
- Alexander, J. C. (2001). *Soziale Differenzierung und kultureller Wandel: Essays zur neofunktionalistischen Gesellschaftstheorie*. Campus-Verlag.
- Ansell, C., & Torfing, J. (Eds.). (2016). *Handbook on theories of governance*. Edward Elgar Publishing.
- Armington, K. (2016). Political institutions. In H. Keman & J. J. Woldendorp (Eds.), *Handbook of research methods and applications in political science* (pp. 234–247). Edward Elgar.
- Barker, R. M. (2010). *Corporate governance, competition, and political parties: Explaining corporate governance change in Europe*. Oxford University Press.
- Bertelli, A. M. (2012). *The political economy of public sector governance*. Cambridge University Press.
- Brante, T. (1988). Sociological approaches to the professions. *Acta Sociologica*, 31(2), 119–142.
- Bulmer-Thomas, V. (2003). *The economic history of Latin America since independence*. Cambridge University Press.
- Carman, J. C. (2009). Nonprofits, funders, and evaluation. Accountability in action. *The American Review of Public Administration*, 39(4), 374–390.
- Chan, A. S., & Fisher, D. (Eds.). (2008). *The exchange university: Corporatisation of academic culture*. UBS-Press.
- Christensen, T., & Laegreid, P. (Eds.). (2007). *Transcending new public management: The transformation of public sector reforms*. Ashgate.
- Coase, R. H. (1937). The Nature of the Firm. *Economica*, 4(16), 386–405.
- Cohen, J. L., & Arato, A. (1994). *Civil society and political theory*. MIT Press.
- Colbeck, C. L. (2008). Professional identity development theory and doctoral education. *New Directions for Teaching and Learning*, 113, 9–16.
- Cutt, J., & Murray, V. (2000). *Accountability and effectiveness evaluation in Non-Profit Organisations*. Routledge.
- Ebener, A. (2008). Institutional evolution and the political economy of governance. In Ebener, A., & Beck, N. (Eds.), *The institutions of the market. Organisations, social systems, and governance* (pp. 287–308). Oxford University Press.
- Edwards, M. (2020). *Civil society*. Polity Press (4th edition).
- Esser, F., & Vliegenhart, R. (2017). Comparative research methods. In Matthes, J., Davies, C. S., & Potter, R. F. (Eds.), *The international encyclopedia of communication research methods*. Wiley and Sons Online Library. <https://onlinelibrary.wiley.com/doi/epdf/10.1002/9781118901731.iccm0035>. Accessed on 19 December 2020.

- Flew, T., & Wilson, J. (2012). WikiLeaks and the challenge of the "Fifth Estate." In M. Ricketson (Ed.), *Australian Journalism Today* (pp. 168–181). Palgrave Macmillan.
- Freidson, E. (2001). Professionalism: The third logic. *Polity*.
- Friedman, A., & Phillips, M. (2004). Balancing strategy and accountability: A model for the governance of professional associations. *Nonprofit Management and Leadership*, 15(2), 187–204.
- Furubo, J. E., Rist, R. C., & Sandahl, R. (2002). *International atlas of evaluation*. Transaction Publishers.
- Giddens, A. (1996). *Konsequenzen der Moderne*. Frankfurt: Suhrkamp.
- Goerres, A., Siewert, M. B., & Wagemann, C. (2019). Internationally comparative research designs in the social sciences: Fundamental issues, case selection logics, and research limitations. In Andreß, H. A., Fetschenhauer, D., & Meulemann, H. (Eds.), *Cross national comparative research* (pp. 75–97). Springer (Special issue 59 *Kölner Zeitschrift für Soziologie und Sozialpsychologie*).
- Goodnow, F. J. (2003). *Politics and administration: A study in government*. Routledge.
- Greenwood, R., Suddaby, R., & Hinings, C. R. (2017). Theorizing change: The role of professional associations in the transformation of institutionalised fields. *Academy of Management Journal*, 45(1), 58–80.
- Han, Y. (2019). The impact of accountability deficit on agency performance: Performance-accountability regime. *Public Management Review*, 22(6), 927–948.
- Hantrais, L. (2008). *International comparative research: Theory, methods and practice*. Palgrave Macmillan.
- Howell, J., & Pearce, J. (2001). *Civil society and development: A critical exploration*. Boulder: Lynne Rienner.
- Inglehart, R. (1998). *Modernisierung und Postmodernisierung: Kultureller, wirtschaftlicher und politischer Wandel in 43 Gesellschaften*. Campus-Verlag.
- Ingvarson, L. (1998). Professional development as the pursuit of professional standards: The standards-based professional development system. *Teaching and Teacher Education*, 14(1), 127–140.
- Jacob, S., Speer, S., & Furubo, J. E. (2015). The institutionalisation of evaluation matters: Updating the international atlas of evaluation 10 years later. *Evaluation*, 21(1), 6–31.
- Kalm, S., & Uhlin, A. (2015). *Civil society and the governance of development: Opposing Global Institutions*. Palgrave Macmillan.
- Kapitanova, J. (2013). *Regeln in Sozialen Systemen*. Springer VS.
- Keman, H., & Pennings, P. (2017). Comparative research methods. In Caramani, D. (Ed.), *Comparative politics* (pp. 49–64). Oxford University Press (4th edition).

- Kettl, D. F. (2015). *The transformation of governance: Public administration for the twenty-first century*. JHU Press.
- Kettl, D. F. (2018). *Politics of the administrative process*. Sage/CQPress.
- Larivière, V., Haustein, S., & Mongeon, P. (2015). The oligopoly of academic publishers in the digital era. *PLoS ONE*, 10(6), 1–15.
- Laville, J. L., Young, D. R., & Eynaud, P. (Eds.). (2015). *Civil society, the third sector and social enterprise: Governance and democracy*. Routledge.
- Lerner, D. (1968). Modernisation: Social aspects. In L. David (Ed.), *International encyclopedia of the social sciences* (pp. 386–402). Macmillan.
- Lovan, W. R., Murray, M., & Shaffer, R. (Eds.). (2003). *Participatory governance: Planning, conflict mediation and public decision making in civil society*. Routledge.
- Lowe, D., & Potter, C. (2018). *Understanding legislation: A practical guide to statutory interpretation*. Bloomsbury Publishing.
- Luhmann, N. (1983). *Legitimation durch Verfahren*. Suhrkamp.
- Luhmann, N. (1988). *Soziale Systeme: Grundriß einer allgemeinen Theorie*. Suhrkamp.
- Mahoney, J. (2010). *Colonialism and post-colonial development. Spanish America in comparative perspective*. Cambridge University Press.
- March, J. G., & Olsen, J. P. (1984). The new institutionalism: Organisational factors in political life. *American Political Science Review*, 78(3), 734–749.
- MacDonald, K. M. (1985). Social closure and occupational registration. *Sociology*, 19(4), 541–556.
- MacDonald, K. M. (1995). *The sociology of the professions*. Sage.
- May, E., Shand, D., Mackay, K., Rojas, F., & Saavedra, J. (Eds.). (2006). *Towards the institutionalisation of monitoring and evaluation systems in Latin America and the Caribbean: Proceedings of a World Bank/Inter-American Development Bank Conference*. IADB/The World Bank.
- Mayntz, R. (2003). From government to governance: Political steering in modern societies. *Summer academy on IPP*, 7–11. [https://www.ioew.de/fileadmin/user\\_upload/DOKUMENTE/Veranstaltungen/2003/CVMayntz.pdf](https://www.ioew.de/fileadmin/user_upload/DOKUMENTE/Veranstaltungen/2003/CVMayntz.pdf). Accessed on 06 June 2019.
- McKeown, A. (2004). Global migration, 1846–1940. *Journal of World History*, 15(2), 155–189.
- Meyer, W., Stockmann, R., & Taube, L. (2020). The Institutionalisation of Evaluation Theoretical Background, Analytical Concept and Methods. In Stockmann, R., Meyer, W., & Taube, L. (Eds.), *Institutionalisation of evaluation in Europe* (pp. 3–34). Palgrave Macmillan.
- OECD (1997). *Promoting the use of programme evaluation*. Public Management Service, OECD (PUMA/PAC) 97(3). October 1997.
- Ojo, A., & Mellouli, S. (2018). Deploying governance networks for societal challenges. *Government Information Quarterly*, 35(4), 106–112.



- Parsons, T. (1939). The professions and social structure. *Social Force*, 17(4), 457–467.
- Parsons, T. (1971). Evolutionäre Universalien in der Gesellschaft. In Zapf, W. (Eds.), *Theorien des sozialen Wandels* (pp. 55–74). Kiepenheuer & Witsch.
- Payne, A., & Phillips, N. (Eds.). (2014). *Handbook of the international political economy of governance*. Edward Elgar Publishing.
- Pérez-Yarahuán, G., & Maldonado, C. (Eds.). (2020). *National monitoring and evaluation systems. Experiences from Latin America*. CLEAR (published in Spanish 2015).
- Pérez-Yarahuán, G., & Maldonado, C. (2020a). Introduction. Reasons for documenting the advances, challenges, and difficulties facing Latin America's monitoring and evaluation systems. In Pérez-Yarahuán, G., & Maldonado, C. (Eds.), *National monitoring and evaluation systems. Experiences from Latin America* (pp. 15–26). CLEAR.
- Pérez-Yarahuán, G., & Maldonado, C. (2020b). Conclusions: Development, innovation and practice. The recent institutionalisation of monitoring and evaluation systems in Latin America. In Pérez-Yarahuán, G., & Maldonado, C. (Eds.), *National monitoring and evaluation systems. Experiences from Latin America* (pp. 380–406). CLEAR.
- Peters, B. G. (2019). *Institutional theory in political science: The new institutionalism*. Bloomsbury Publishing USA.
- Pollack, D. (2016). Modernisierungstheorie – Revised: Entwurf einer Theorie moderner Gesellschaften. *Zeitschrift Für Soziologie*, 45(4), 219–240.
- Richardson, A. J. (1997). Social closure in dynamic markets: The incomplete professional project in accountancy. *Critical Perspectives on Accounting*, 8(6), 635–653.
- Rosenbaum, A. (2006). Cooperative service delivery: The dynamics of public sector-private sector-civil society collaboration. *International Review of Administrative Sciences*, 72(1), 43–56.
- Rosenstein, B. (2013, December). Mapping the status of national evaluation policies. *Parliamentarians Forum on Development Evaluation in South Asia and EvalPartners*. <http://www.pfde.net/index.php/publications-resources/global-mapping-report-2015>. Accessed on 14 October 2020.
- Rüschemeyer, D. (1983). Professional autonomy and the social control of expertise. In Dingwall, R. & Lewis, P. (Eds.), *The sociology of the professions. Lawyers, doctors and others* (pp. 38–58). MacMillan St. Martin's Press.
- Segovia, D. (2013). Latin America and the Caribbean: Between the OAS and CELAC. *European Review of Latin America and Caribbean Studies*, 95, 97–107.
- Schiele, B., Claessens, M., & Shi, S. (Eds.). (2012). *Science communication in the world: Practices, theories and trends*. Springer Science & Business Media.
- Schimank, U. (1996). *Theorien gesellschaftlicher Differenzierung*. Springer-Verlag.

- Schimank, U. (2001). Teilsysteminterdependenzen und Inklusionsverhältnisse. Ein differenzierungstheoretisches Forschungsprogramm zur System- und Sozialintegration moderner Gesellschaft. In Barlösius, E., Müller, H., & Sigmund, S. (Eds.), *Gesellschaftsbilder im Umbruch. Soziologische Perspektiven in Deutschland* (pp. 109–130). Springer-Verlag.
- Schofield, N., & Caballero, G. (Eds.). (2015). *The political economy of governance: Institutions, political performance and elections*. Springer.
- Schultz, J. (1998). *Reviving the Fourth Estate. Democracy, Accountability and the Media*. Cambridge University Press.
- Stachura, M., Bienfait, A., Albert, G., & Sigmund, S. (Eds.). (2009). *Der Sinn der Institutionen: Mehr-Ebenen-und Mehr-Seiten-Analyse*. Springer-Verlag.
- Stockmann, R. (2008). *Evaluation and quality development: Principles of impact-based quality management*. Peter Lang.
- Stockmann, R. (2013). Evaluation in der Entwicklungszusammenarbeit. In J. Wilhelm & H. Ihne (Eds.), *Einführung in die Entwicklungspolitik* (3rd ed., pp. 541–562). LIT Verlag.
- Stockmann, R., & Meyer, W. (Eds.). (2016). *The Future of Evaluation: Global Trends, New Challenges, Shared Perspectives*. Palgrave Macmillan.
- Stockmann, R., Meyer, W., & Taube, L. (Eds.). (2020). *Institutionalisation of evaluation in Europe*. Palgrave Macmillan.
- Tocqueville, A. (2003). *Democracy in America and two essays on America*. Penquin.
- Traugott, M. (2013). *Emile Durkheim on institutional analysis*. Chicago University Press.
- Wong, J. (2014). Comparing Beyond Europe and North America. In Engeli, I., & Rothmayr, A. C. (Eds.), *Comparative Policy Studies. Conceptual and Methodological Challenges* (pp. 163–184). Palgrave Macmillan.
- Zapf, W. (Eds.). (1991). Die Modernisierung moderner Gesellschaften: Verhandlungen des 25. Deutschen Soziologentages in Frankfurt am Main 1990 (Vol. 25). Campus Verlag.

PART II

---

National Developments



## Evaluation in Argentina

*Pablo Rodríguez-Bilella and Esteban Tapella*

### GENERAL OVERVIEW

The institutionalisation of evaluation, understood as the incidence and permanent anchorage of formal and informal rules, structures and evaluative processes, has been historically marginal within the framework of priorities set by the different Argentine governments. An indicator thereof is the absence of policies oriented towards the strengthening of both institutional capacities and the integrality of the state evaluation function in the country, in contrast to some other countries of the region. Since there is no unit in charge of explicitly planning, implementing and coordinating evaluation of programmes and policies in the organic structure of the executive branch, those competencies and responsibilities are divided among several organisms (ministries, secretaries or departments). Thus, the *Ministerio de Hacienda* (Ministry of Finance) is in charge of the

---

The information in this chapter was updated until the end of 2019, when there was a change of the central government.

---

P. Rodríguez-Bilella (✉) · E. Tapella  
CONICET / Universidad Nacional de San Juan, San Juan, Argentina

© The Author(s), under exclusive license to Springer Nature  
Switzerland AG 2022

41

R. Stockmann et al. (eds.), *The Institutionalisation of Evaluation in the Americas*,  
[https://doi.org/10.1007/978-3-030-81139-6\\_2](https://doi.org/10.1007/978-3-030-81139-6_2)

physical-financial tracking of budget programmes; the *Jefatura de Gabinete de Ministros* (Chief of the Cabinet of Ministers) is responsible for monitoring and evaluating programmes and policies; and diverse sectorial ministries are in charge of defining the needs and characteristics of their evaluations according to the will or interest of their politic heads or, what is often more frequent, of particular international requirements.<sup>1</sup> In September 2018, eight ministries were subdued to secretaries, reducing the salience of their policies and that of the corresponding evaluation efforts.

Despite some progress made towards the institutionalisation of evaluation systems (which has been reached throughout the consideration of the latter in strategic plans), the importance given to the so-called neo-universal policies<sup>2</sup> and their presence in flow charts concerning diverse areas, evaluation experiences within the distinct spheres of the executive branch form fragmented and disjointed experiences (Canievsky, 2007). Therefore, they do not account for a global and integral strategic vision, capable of favouring and boosting both the promotion and use and the building of evaluation capacities.

Decree No. 292 of 2018, titled '*Lineamientos de Monitoreo y Evaluación*' (Monitoring and Evaluation Guidelines), became the regulation establishing the technical basis for constant monitoring and evaluation of policies, programmes, plans and projects with social impact implemented by ministries and national organisms. According to such decree, the *Consejo Nacional de Coordinación de Políticas Sociales* (National Council for the Coordination of Social Policies) is entitled to develop each year the *Plan Anual de Monitoreo y Evaluación de Políticas y Programas Sociales* (Annual Plan for Monitoring and Evaluation of Social Policies and Programmes), whose implementation will be mandatory for all public organisms and entities carrying out social policies, programmes, plans and projects financed by the *Tesoro Nacional* (National Treasury) or international organisms.

<sup>1</sup> Referring to the demands or requests made by programmes funded by foreign agencies and evaluated by them.

<sup>2</sup> Neo universal policies are those policies aiming at overcoming the neoliberal approach of public policies, where the social aspect plays a primordial role given the absence or previous localisation of state attention. In discourse, they contemplate issues related to development concerning social inclusion and income distribution. The Universal Child Allowance for Social Protection is usually quoted as an example thereof in the Argentinian case.

Civil society organisations generally maintain a very close connection with the practice of evaluation. While they focus on the implementation of several initiatives (concerning social development, lobbying, incidence, etc.), they can receive or conduct any type of evaluation, fundamentally drawing on external demands coming from an agent funding their projects.

## INSTITUTIONAL STRUCTURES AND PROCESSES (POLITICAL SYSTEM)

### *Evaluation Regulations*

Even though at State level there are no laws establishing an integral evaluation system focusing on the actions carried out by the national public administration, since 1990, distinct aspects are indeed regulated, concerning monitoring and evaluation fundamentally centred in the managing for results perspective, evaluation of performance and monitoring of expenditure (Aquilino & Amaya, 2015). In the Argentinian case, characterised by a strong presidential and representative democracy system and a two-chamber parliament, this regulation refers to the legal regulations established by the executive branch, oriented towards the implementation and completion of existent laws, generally binding.

In 2013, the *Jefatura de Gabinete de Ministros* (Chief of the Cabinet of Ministers) stipulated the creation of the *Programa de Evaluación de Políticas Públicas* (Public Policies Evaluation Programme), whose main aim was the institutionalisation of evaluation processes concerning public policies within the public administration and the maximisation of capacities regarding their development. There is no evidence about this programme being used as an instrument for improvement of policies or their management, even though training instances for the public administration personnel were generated; a database of technicians and public servants in charge of monitoring and evaluation was developed; a handbook on the evaluation of public policies was published; and an evaluation pool was created.

The *Jefatura de Gabinete de Ministros* (Chief of the Cabinet of Ministers) emerged as an authority concerning the natural application of an integral evaluation system, since it is not only in charge of state administration but it was also assigned specific competencies related to the

coordination of evaluation in the national state. However, one thing is certain, its historical political weakness has prevented it to evolve as such.

### *Laws of Sectoral Regulations*

The notion of sector speaks of different political fields, such as health, education, infrastructure, which can be understood as subsystems of the political or social system. In this respect, Argentine laws reflect the absence of an integral system concerning the monitoring and evaluation of public policies, even though some of them constitute aspects of a potential national system of evaluation, regulating different subjects concerning monitoring and evaluation (Aquilino et al., 2017).

Even if the performance audit concerns the objective and systematic exam of effectiveness (achievement of goals) and efficiency (economic use of resources), based on external criteria, it is relevant to examine those regulations transcending this focus and having evaluative components. For instance, the *Ley N° 24.156 de Administración Financiera y Sistemas de Control* (Law No. 24.156 of Financial Management and Control Systems) passed in 1992 established that the investment account to be annually presented to the congress should include comments on the degree of accomplishment of the goals and objectives estimated in the budget, as well as on the behaviour of costs and efficiency indicators regarding public production and financial management of the national public sector.

The *Dirección Nacional de Inversión Pública del Ministerio de Economía y Finanzas Públicas* (National Bureau for Public Investment of the Ministry of Economy and Public Finances), as the responsible body of the system, develops pertinent indicators and decision criteria to formulate and evaluate public investment programmes and projects. For its part, the *Ley N° 25.152 de Solvencia Fiscal y Calidad de la Gestión Pública Nacional* (Law N° 25.152 of Fiscal Solvency and Public Management Quality), passed in 1999, established the Programme for Evaluation of Public Expenditure with the purpose of increasing the quality of services through the systematic evaluation of costs related to results. Finally, the *Auditoría General de la Nación* (General Audit of the State) is the public sector's external control body, which depends on the National Congress.

Stated briefly, regulations supporting the organisational aspects of control, monitoring and evaluation in public administration agencies show an old and closed legislation, which focuses on the evaluation of expenditure, and thus is not capable of generating an integral approach

enabling the analysis and assessment of the impact of State action in the daily life of the inhabitants of the country. The evaluative component is too weak to interpret the data produced, thus emphasising only one perspective of monitoring. Evaluation experiences oriented towards learning, allowing to draw from the lessons of past interventions in order to improve plans and programmes in the future are even less shown. More recent law proposals turn out instrumental and lack a long-term vision.

*Policies or Strategies Regarding Evaluation and Its Use*

The notion of ‘politics’ refers to a coherent strategy of action focused on a specific subject. If understood broadly, as a group of issues reflecting a national strategy, it might seem that there are no policies or consolidated strategies regarding evaluation at national level in Argentina. Nonetheless, it is possible to identify a series of actions at sectoral level, which are reflected in the existence of monitoring and evaluation units among different ministries.

Towards the end of 2015, the *Dirección de Gestión y Evaluación de Programas del Ministerio de Salud y Desarrollo Social* (Bureau of Management and Evaluation of Programmes from the Ministry of Health and Social Development) was responsible for synthetising the information required for territorial interventions. As of the management change, new regulations granted the *Sistema de Información, Evaluación y Monitoreo de Programas Sociales* (SIEMPRO, System of Information, Evaluation and Monitoring of Social Programmes, according to the Spanish acronym) the capacity to collect and integrate information about the programmes of the ministry, mainly information about the beneficiaries (Canievsky, 2007). SIEMPRO depends on the *Consejo Nacional de Coordinación de Políticas Sociales* (National Council for the Coordination of Social Policies), a space where state areas implementing social policies are articulated, in order to achieve a correct and more efficient management of resources. Despite counting on a specific area for the development of evaluations, SIEMPRO’s actions have been delimited in this last decade, as of the development of evaluation areas in the different ministries.

The *Ministerio de Hacienda* (Ministry of Finance) has developed methodologies and criteria for the formulation and evaluation of investment projects for their homogenous use in the implementing agencies of the national public administration. The *Oficina Nacional de Presupuesto* (National Budget Office) intervenes in the formulation, programming



of execution, modifications and evaluation of the public administration budgets.

The *Ministerio de Educación, Cultura, Ciencia y Tecnología* (Ministry of Education, Culture, Science and Technology) considers both learning (school performance) and policies and institutions are subject to evaluation, promoting the information system and evaluation of education as a tool for improvement. The evaluation of the education system is the responsibility of the *Secretaría de Evaluación Educativa* (Secretariat of Educational Evaluation), in which the *Red de Evaluación Federal para la Calidad y Equidad Educativa* (Federal Evaluation Network for Educational Quality and Equity) is located. This network aims to establish an articulated operation between the areas with competence in evaluation in each jurisdiction and the *Ministerio de Educación, Cultura, Ciencia y Tecnología* (Ministry of Education, Culture, Science and Technology), in order to facilitate the implementation of a national evaluation system and promote improvements in quality and equity. For its part, the *Secretaría de Políticas Universitarias* (Secretariat of University Policies) is oriented to the development and evaluation of plans, programmes and projects for the development of the university higher education system. The *Comisión Nacional de Evaluación y Acreditación Universitaria* (CONEAU, National Commission for University Evaluation and Accreditation) is also part of this ministry. It was created with the mission of ensuring and improving the quality of university careers and institutions through accrediting activities and evaluations of graduate and undergraduate careers, evaluation of institutional projects regarding new establishments, as well as the monitoring of university institutions. The structure of this ministry also includes the *Agencia Nacional de Promoción Científica* (National Agency for Scientific and Technological Promotion), which relies on the *Unidad de Evaluación y Aseguramiento de la Calidad* (Unity of Evaluation and Quality Assurance) for the monitoring, evaluation and assurance of the quality of the activities developed for the promotion of science, technology and productive innovation.

The *Ministerio de Salud* (Ministry of Health) became the Secretariat of the *Ministerio de Salud y Desarrollo Social* (Ministry of Health and Social Development) in September 2018. It houses the *Dirección de Gestión y Monitoreo de Programas y Proyectos Sectoriales y Especiales* (Directorate of Management and Monitoring of Sectoral and Special Programmes

and Projects), whose primary responsibility is to coordinate and implement a comprehensive system regarding the supervision, monitoring and evaluation of the management of health facilities.

In the *Ministerio de Producción y Trabajo* (Ministry of Production and Labor), the *Subsecretaría de Programación Técnica y Estudios Laborales* (Under secretariat of Technical Programming and Labor Studies) was dissolved in 2018. This entity was responsible for evaluating the impact of policies, plans, programmes and projects in the labour market, labour relations, the income of workers and the social security system. Within the framework of the reduction of state expenses, the recruitment of professionals with extensive experience was discontinued and various professionals were relegated to non-substantive tasks.

The *Ministerio de Ciencia, Tecnología e Innovación Productiva* (Ministry of Science, Technology and Productive Innovation), dissolved as such, became the *Secretaría del Ministerio de Educación, Cultura, Ciencia y Tecnología* (Secretariat of the Ministry of Education, Culture, Science and Technology) in September 2018. It undertakes the responsibility of designing and supervising the application of criteria and procedures for the evaluation of science, technology and innovation organisations. In the period between 2010 and 2015, about 20 programmes were evaluated ex ante, mid-term and ex post by the *Programa Nacional de Evaluación y Fortalecimiento de Programas de Ciencia, Tecnología e Innovación* (PRONEP, National Programme for Evaluation and Strengthening of Science, Technology and Innovation Programmes). After the management change in 2015, PRONEP received no demands for further evaluations.

The *Ministerio de Agroindustria* (Ministry of Agribusiness) became a secretariat in September 2018, depending since then on the *Ministerio de Producción y Trabajo* (Ministry of Production and Labor). Although it does not have a central area oriented towards monitoring and evaluation, there are administrative units carrying out these functions in the decentralised agencies which depend on it. These include the work of the *Instituto Nacional de Tecnología Agropecuaria* (National Institute of Agricultural Technology), which relies upon the *Dirección de Planificación, Monitoreo y Evaluación* (Directorate of Planning, Monitoring and Evaluation). This entity performs, among other functions, ex ante, mid-term and ex post evaluations of the projects.

The outlined sectoral experiences point out the absence of a strategic framework clearly defining their function and objectives within an integral

evaluation system. The latter reveals weaknesses concerning their institutional insertion, disarticulation and lack of coordination with each other (Acuña et al., 2016). They have developed in diverse historical contexts and have arisen in the face of specific needs and situations, without having a global and integral vision that gives rise to a policy of capacity building, promotion and use of evaluation.

*Content of Administrative Strategies or Regulations Regarding Independence, Quality, Impact Orientation and Budget Available for Evaluation*

The evaluation of the results and impacts of public policies in Argentina is based upon the intention and will of officials, building up a scenario of low relevance regarding their demand and use by decision makers. Different ministries and public departments have specific areas, which are oriented towards the evaluation of policies and programmes, even though their existence is not a rule for all sectoral instances. In some provinces, monitoring and evaluation systems were developed for specific plans or programmes, especially in the social sector as in the case of Santa Fe. In the same province, the city of Rafaela generated an evaluation agency at municipal level. However, the different experiences and initiatives are disjointed and only a few strategic plans and programmes are rigorously analysed in order to understand whether their effects are desired or to what extent they contribute to achieving the conceived social change.

*Evaluation and Its Use in Connection with Parliamentary Structures*

The connection existing between the Argentine parliament and the practices of monitoring and evaluation of public policies is minimal. There is no regular practice regarding the report of the results generated by the evaluation of programmes and public policies.

*The Practice of Evaluation*

It should be noted that ex-ante, concomitant and ex-post evaluations are carried out in the field of national public administration, particularly evaluations of results. As for instruments, surveys, interviews, administrative databases as well as the emphasis on management data with ad hoc developed tools are mostly used, usually involving quantitative, qualitative and mixed elements.

Partial breakthroughs achieved have not been able to realise the necessary link between knowledge about the achievements or failures regarding results (and their reasons) and decision-making (improvement in management, changes of course, budgetary reformulations, transparency of management, etc.) (Neirotti, 2000, 2001). Along with this, the unequal degree of evaluability of plans and programmes points out the importance of strengthening the institutional context, the quality of public policy design as well as the development of capacities within implementing agencies (Aquilino et al., 2015).

#### *Relationship Between Internal and External Evaluations*

While external evaluation refers to evaluations carried out by people (experts) not being part of the implementation or funding organisation, internal evaluation refers to evaluations carried out by the same organisation implementing a programme. In turn, self-evaluation is a specific form of internal evaluation, in which people implementing the programme are also in charge of evaluating it (Stockmann & Meyer, 2016).

In the case of Argentina and in the context of national public administration, organisational forms concerning evaluations have emphasised on internal evaluations, with an important focus on the evaluation of expenditure and with much less emphasis and ability to address public policy interventions more comprehensively, as well as to learn from them in order to redress future plans and programmes. Although different credit organisations promoted the development of evaluations carried out with international funding, in many cases evaluations continued to be public and internal (Vinocur & Halperin, 2004).

#### *Relationship Between Process and Impact Evaluations*

Process evaluations focus on the planning stage and on the implementation activities of the intervention, particularly on its inputs and products. For their part, impact evaluations focus on the actual results and impact of the measures of the programme, both expected and desired as well as those unexpected and undesirable. Both rely on the availability of reliable data on the characteristics of the processes, inputs used and costs, as well as on the progress made in relation to the goals established concerning products and results. That is why a fundamental component in the development of an evaluation system is connected to the strengthening and legitimisation of information systems.

In this sense, regarding state institutionalisation, Argentina faces a serious deficit in terms of capacities linked to planning and generation of information. Existing diagnoses account for the lack of national mid and long-term plans in which the objectives, goals and results of different policies, programmes and sectors are articulated in a consistent manner. Argentina has regressed in terms of the quantity and quality of data and trust in official sources has been severely broken, which directly affects the very quality of its democracy.

### *Independent Evaluation Institute*

In Argentina, there is no independent institute or evaluation unit within the organisational structure of the national executive branch explicitly in charge of the planning, implementation and coordination of programme and policy evaluation. Therefore, these competencies and responsibilities are segmented into various agencies and differ markedly between ministries, government agencies, provinces and municipalities.

This reality is particularly critical in the Argentinian case, a federal state, in which a national state, provincial states and the state of the Autonomous City of Buenos Aires coexist, each one made up of the corresponding administrations. Along with these, municipal governments depending on provinces, which have a certain degree of autonomy, should also be considered.

Five bills have been presented in the national congress, in order to create the *Agencia Nacional de Evaluación de Políticas Públicas* (National Agency for the Evaluation of Public Policies), but the legislative branch has not dealt with any of them. This legal vacuum reveals a lack of consensus among the different political forces on the relevance that evaluation can have for the improvement of public policies in the short, medium and long terms.

## *Use of Evaluations*

### *Sectors of Good Performance Regarding Evaluation and Use*

In Argentina, information generated by evaluations is mostly used for accountability and budget allocation. The *Ministerio de Hacienda* (Ministry of Finance) has a long track record concerning the development of methodologies and the conduction of evaluations of investment projects. The *Oficina Nacional del Presupuesto* (National Budget Office),

located inside it, plays a significant role at all stages regarding the budget, particularly its monitoring.

Various evaluations are carried out at different levels by the *Secretaría de Evaluación Educativa del Ministerio de Educación, Cultura, Ciencia y Tecnología* (Secretariat of Educational Evaluation of the Ministry of Education, Culture, Science and Technology). Learning-oriented use and redesign are usually based upon the definition of institutional commitments, which are followed up later. However, there are limitations concerning its use, due to the lack of useful information for managers. In the case of CONEAU, its regular evaluations and the encouragement of self-evaluation regarding postgraduate training instances position it as a sector increasingly reinforcing its performance in the field of evaluation.

The *Secretaría de Ciencia, Tecnología e Innovación Productiva* (Secretariat of Science, Technology and Productive Innovation) is responsible for the design and supervision of the application of criteria and procedures within the evaluation of science, technology and innovation organisations. In the case of PRONEP, more than one-third of the evaluations are estimated to have been used in the reformulation of work lines and budget discussion.

For its part, the use of debate-oriented evaluations within the scope of the parliament is also limited by the multiplicity and complexity regarding information provided in evaluation reports. While there are also some routine mechanisms that require the executive branch to provide information on government plans and their measures concerning the legislative branch, there are deficits regarding the horizontal accountability process carried out between the executive and the legislative branch (Aquilino et al., 2016). Some of them refer to the poor quality of information provided by the Chief of Staff and the president in their speeches and the very few responses received by the Legislative branch concerning requests sent to the Executive branch.

#### *Sectors of Limited Performance Regarding Evaluation and Use*

The *Consejo Nacional de Coordinación de Políticas Sociales* (National Council for the Coordination of Social Policies) was established in order to articulate the areas of the national state implementing social policies, with the purpose of improving their impact and efficiency. Its privileged instrument for the development of evaluations has been SIEMPRO, which has limited itself to supervise evaluations carried out at the request of multilateral organisations. SIEMPRO shows very weak institutional

consolidation, which raises questions about its operation and usefulness (Neirotti et al., 2015).

The *Secretaría de Salud* (Ministry of Health), under the *Ministerio de Salud y Desarrollo Social* (Ministry of Health and Social Development), relies on regulations aiming at the coordination and implementation of a comprehensive system of supervision, monitoring and evaluation of the management of health facilities. The absence of systematic monitoring mechanisms regarding the use of evaluations carried out, as well as the scarce presence of the latter in the information sector sets up a scenario where they account for very little impact.

The *Secretaría de Agroindustria* (Secretariat of Agribusiness) relies on different administrative units complying with monitoring and evaluation functions in the decentralised agencies that depend on it. Although they carry out the ex-ante, mid-term and ex-post evaluations of different types of projects, a greater limitation in the performance and use of evaluations, where the relevance, importance and characteristics of the latter depend closely on the will and interest of its policy-makers, with relative independence from existing regulations, is to be seen. Frequently, it is the international requirements that trigger the need for evaluations regarding programmes financed by these instances.

#### *Aspects on Which Evaluations Focus*

Progress made in the creation of evaluation structures in Argentina formed the basis for the development of monitoring and evaluation functions in the country. However, a greater place for evaluation within the national public administration did not result from a planning instance homogenising and guiding these efforts. Thus, the connection existing between knowledge generated about achievements or failures concerning the results, along with their reasons, and the necessary decision-making resulting from these findings was weak.

This allows to realise that a recurring aspect in the functionality of sector evaluations concerns fiscal control and legitimisation of the intervention. The components oriented to indicate improvements in the section or changes of course in it are much more limited. This emphasis on management responsibility and transparency sharply emphasises the accountability of learning instances in evaluations.

### *Actors Who Make Regular Use of Evaluation and Its Results*

Public bodies in Argentina have privileged work instances on monitoring rather than the evaluation of programmes and policies, thus showing the predominance of a budgetary and operational perspective of evaluation with a focus on spending control and performance indicators (product, coverage, processes and, to a lesser extent, results). Despite certain progress and several attempts to systematically integrate monitoring and evaluation into public policies, a general trend has remained, characterising the practice of evaluation as a timely and isolated instance, which is disconnected from its development within the framework of a greater cohesive system.

Different evaluation systems tend to use information produced by themselves, with a low level of exchange. This happens because the decision of producing and disseminating such information lies within the discretion of the body that generated it, revealing the weakness of formal mechanisms concerning the relationship among the different instances of government, in order to learn about what has been done and the difficulties and lessons learned.

At the level of the provinces constituting the Argentine republic, the picture is heterogeneous in terms of capacities, legal frameworks and institutionalisation of evaluation. While some provinces have made progress in adopting mid- and long-term strategic government plans (which are key to the evaluability of public policy results), others have managed to develop skills regarding managing for results as a relevant step towards a systematic use of evaluation.

### *Quality of Evaluations*

There are no meta-evaluation mechanisms enabling control or monitoring of the quality of evaluations. Their quality is only reviewed by those who contract them, thus their relevance, rigour, quality and their consequent use remain private. The absence of mechanisms capable of promoting the quality of evaluation, such as the use of evaluation standards, as well as the absence of ethical codes or guiding principles for evaluators, contributes to the impossibility to appreciate this dimension.

### *As a Closing Section*

Sectoral evaluation experiences in Argentina do not rely on a strategic framework integrating monitoring and evaluation functions into the daily management of the public administration, and clearly stipulating their



function and objectives within a comprehensive evaluation system. Incipient proposals aiming at some progress on the integration of the different evaluation spaces in different national ministries as well as the extension of evaluation practices stopped with the shift regarding government management in 2015.

Although there has been a revaluation of evaluation and the relationship between evaluation and the improvement of quality regarding policies is now more explicit, Argentina does not rely on an articulated monitoring and evaluation system. It rather has some agencies or units that carry out policy evaluations with almost no articulation with each other. This segmentation of the different evaluation structures contributes to the fact that evaluation has not been effectively reflected in programme and policy implementation budgets. In general, connections between evaluations and strategic policy levels are scarce. There is also a very limited capacity concerning appropriation of results by actors participating in programmes or policies. The latter constitutes a very relevant weakness of the Argentinian state, unable to seize the strategic importance of evaluation along state reform processes as an instance capable of enriching the public debate and promoting collective learning about initiatives, strategies and actions aiming at influencing the functioning of the State.

Specific initiatives aiming at the generation of changes and progress in terms of evaluation in the context of improving state management have not managed to permeate bureaucratic structures in national, provincial or municipal instances. These efforts did not adequately consider the administrative institutional context in which they would operate, characterised by the predominance of delegative democracy practices that reinforce the political dependence of the implementing, regulating and controlling bodies with respect to the public policy-making bodies (Zaltsman, 2006).

Without a national policy or a governing agency or unit, Argentina reduces the ability to learn about the results of state interventions because of the lack of regulatory frameworks and evaluative practices enabling and requiring the measurement and systematic analysis of plans and programmes. The absence of appropriate legislation and institutional mechanisms guaranteeing the monitoring and evaluation of public policies characterises the institutionalisation of evaluation as low, fragmented and isolated—both as a practice and as a system.

## DIFFUSION IN SOCIETY AND ACCEPTATION (SOCIAL SYSTEM)

### *Institutionalised Use of Evaluations by Civil Society*

The notion of civil society refers to self-organisation and self-management carried out by citizens, especially in the form of associations, non-governmental organisations (NGOs) or non-profit organisations. Within the national context, independent evaluators or independent evaluation firms have evaluated small and mid-scale projects carried out or financed by civil society instances, such as NGOs, foundations, corporate social responsibility (CSR) programmes.

In Argentina, evaluation firms exceptionally focus only on evaluation, since the national evaluation market is limited. These evaluations are often carried out by academics or university research teams, which submit proposals to the calls for evaluation of these projects. Even though evaluation approaches are usually pluralistic, they tend to include mixed methods contemplating documentary revision, interviews, observation and visits to the territory where the project is implemented. A growing interest in data triangulation is observable, whereas Terms of Reference (ToR) usually demand evaluations to consider Organization for Economic Cooperation and Development/Development Assistance Committee criteria regarding relevance, effectiveness, efficiency, sustainability and impact.

In those cases where civil society has played a significant role implementing the intervention, evaluations usually focus on the performance of the programme or project, assessing results and drawing lessons that allow the improvement of the NGOs or funding agency's general programming. Sometimes, these evaluation practices are carried out by inner members of the organisation, based upon their knowledge management area.

Another aspect related to the use of evaluation within civil society is the publication of a handbook on the 'Monitoring and Evaluation of Private Social Investment' (Potenza, 2018), in 2018, by '*Grupo de Fundaciones y Empresas*'. The latter is a non-profit civil association gathering foundations and enterprises committed to sustainable development, which increasingly acknowledge the importance of the evaluation of programmes and projects managed by corporations and corporate foundations. Prominent among the factors explaining the aforementioned importance are the need for capitalising lessons learned from experience

as well as the requirement for accountability on the use of the resources funding private initiatives and the results obtained.

### *Public Perception and Discussion on Evaluation and Evaluation Results*

The dissemination and publicity of evaluations contribute to the improvement of the State's control system, to the perception society has about the responsibility of decision agents and to the inputs for policy results as well. This is why transparency processes are relevant for the reinforcement of the institutionalisation of evaluation in the country, inasmuch access to information constitutes an essential requirement in order to develop evaluation processes and enable the participation of civil society in public policy monitoring and evaluation.

In Argentina, as pointed out in previous sections, evaluation experiences carried out within the national public administration scope are not usually socialised in a broad manner.

Decree No. 1.172 regarding access to public information, issued in 2003, aims to guarantee the principle of public access to official documents—concerning government actions—and the right of accessing public information. Its objective was to strengthen the relationship between the state and civil society, favouring the perception of the latter about the degree of effectiveness and efficiency of public administrations, arbitrating therefore transparency processes.

Along with this, the presence of civil society organisations in instances close to the monitoring and evaluation of government actions has been minimal and marginal, in a context characterised by the narrow space allocated to the promotion and study of improvements, strengthening and state innovation. If this is the reality in the space of government actions, where evaluation experiences are both scarce and poorly disseminated and socialised, the outlook concerning the same civil society organisations becomes acute.

The foundations (e.g. *Fundación Poder Ciudadano*—linked to Transparency International—*Fundación Nueva Generación Argentina*, *Fundación Plus*, *Fundación Conciencia*, etc.) show a minimal and marginal interest for evaluation and concentrate their action on citizen training on state control bodies (their functions, reporting mechanisms, the information they produce, etc.), the formation of electoral observatories, construction of indexes regarding the perception of corruption,

etc. The public debate is also limited in this regard and it is far from constituting a mechanism of control, even an indirect one.

### *Demand for Evaluation Within Civil Society*

In the Argentinian context, no significant interest concerning citizen participation in public policy evaluation processes is to be seen. Although there have been some specific experiences regarding advisory councils within the framework of some social programmes, these have been characterised by their weakness in the development of their functions.

An example of these advisory councils within the framework of government programmes was the *Jefes y Jefas de Hogar* (Heads of Household, both male and female) programme, with a proposal for decentralised and participatory management via such councils—at the national, provincial and municipal level—composed of representatives of NGOs, faith-based institutions, workers ‘and employers’ organisations and by government officials, with the main function of controlling both government actions and the execution of public funds. Another example is the *Consejo Nacional de Coordinación de Políticas Sociales* (National Council for the Coordination of Social Policies), which has a *Consejo Consultivo Nacional de Políticas Sociales* (National Consultative Council for Social Policies). This council is composed of government and business representatives, labour union organisations, social organisations and faith-based institutions, and it is authorised to promote proposals tending to improve and facilitate the territorial articulation of social plans, among others.

In these and other similar cases, advisory councils have often tended to constitute merely formal figures without a concern for the evaluation of the actions faced. In general, approaches have been limited to performance reports, paying attention to inputs delivered and activities deployed, without moving towards more relevant analysis on the outcomes and impact of the interventions.

Thus, while citizen participation regarding public policy evaluation processes would contribute both to achieve greater transparency regarding public management and to strengthen accountability practices by those responsible for executing programmes and policies, a manifest interest in the involvement of the population throughout evaluation in the processes of policy improvement is not detected. A factor possibly explaining this reality is the absence of greater institutionalisation of evaluation (also occasionally understood as evaluation culture). It has not been

possible to disconnect its generation of value judgements from the idea of recrimination or sanction as the ultimate goal, thus preventing a deeper understanding of public interventions aimed at its improvement and, therefore, that of evaluation as an instrument for such transformation. Without a change in this perception, it will not be possible to positively integrate evaluation into the representation of civil society, generating most of all discomfort (due to its consumption of time and energy), distrust (due to its emphasis on administrative control) and disinterest (since its potentialities are not visualised).

An initiative aiming directly at the institutionalisation of evaluation in the country, whose consolidation is still ongoing, is the *Red EvaluAR* (Evaluar Network), the Argentine Evaluation Network (Ortale, 2015). With more than fifteen years of existence, it took renewed momentum from 2013 and has brought together some 200 professionals from all over the country, involved and interested in the field of evaluation. The network aims at the promotion of exchange, dissemination, formation and development of evaluation capacities among professionals and interested groups. At the regional level, it is part of *Red de Seguimiento, Evaluación y Sistematización en América Latina* (ReLAC, Monitoring, Evaluation and Systematisation Network in Latin America), a network of networks which aims both at contributing to the reinforcement of monitoring and evaluation capacities and professionalising the role of evaluation in Latin America and the Caribbean. On the international front, it is part of the International Organisation for Cooperation in Evaluation. In recent years, it has developed training sessions and research projects on evaluation in the country, prominent among these is the ‘*Mapa Diagnóstico de la Evaluación en Argentina*’ (Diagnostic Map of Evaluation in Argentina, Aquilino & Amaya, 2015), to be presented later.

## PROFESSIONALISATION (PROFESSIONALISATION SYSTEM)

### *Academic Courses of Studies, Trainings, Etc.*

In recent years, the offer of training instances regarding evaluation has expanded in Argentina. As for Master’s degrees, it is worth mentioning the *Maestría en Evaluación de Políticas Públicas* (Master in Public Policy Evaluation) in the *Universidad de Entre Ríos*, the *Maestría en Planificación y Evaluación de Políticas Públicas* (Master in Planning and Evaluation of Public Policies) in the *Universidad Nacional de San Martín*

and the *Maestría en Evaluación de Proyectos* (Master in Project Evaluation) in the *Universidad del CEMA*. Since 2016, the *Universidad Nacional de Lanús* developed jointly with the National University Arturo Jauretche a *Especialización en Evaluación de Políticas Públicas* (Specialisation in Public Policy Evaluation). The *Universidad Nacional de San Martín* also has a *Especialización en Evaluación de Políticas Públicas* (specialisation in Public Policy Evaluation). The *Diplomado en Gestión y Control de Políticas Públicas* (Diploma in Management and Control of Public Policies) of FLACSO (Latin American School of Social Sciences), with contents close to evaluation, can be added to these.

A space for training and debate targeting government officials, professionals, teachers and researchers dedicated to monitoring and evaluation processes resulted in the national seminar ‘The evaluation of public policies in the current scenario of transformations in the state’, organised by the *Universidad Nacional de Lanús* in September 2014. An important subsequent publication was made about this seminar (Neirotti, 2015). Giving continuity to this event, at the end of 2017 the same university created the *Agenda Compartida* (Shared Agenda) programme, which has become a space to present different visions about the role of evaluation in the current framework as well as to outline some of the challenges it faces. This programme aims at bringing together actors from academia, national public administration agencies and evaluation professional networks. In August 2017 the University organised the International Seminar: ‘*Evaluación y toma de decisiones. Diálogos entre políticos y académicos para fortalecer la democracia*’ (Evaluation and decision-making. Dialogues between politicians and academics to strengthen democracy), with the participation of academics, evaluators and decision makers.

In recent years, different research programmes have begun to address the field of evaluation as an object of analysis. This is the case of the *Programa de Estudios del Trabajo, el Ambiente y la Sociedad* (Work, Environment and Society Studies Programme) at the *Universidad Nacional de San Juan*, which develops studies around participatory evaluation, use and adoption of the results obtained through monitoring and evaluation processes; including training and transfer actions arising from these studies. Through extension activities, different public sector officials and participants in the Master’s programmes in social policy and territorial development are trained in various methodologies for public policy evaluation, such as the *EvalParticipativa* (Participative Evaluation) initiative,

a recent initiative launched in cooperation with the German Institute of Evaluation for Development Cooperation. Other similar teams at the *Universidad Provincial de Córdoba* and the *Universidad Nacional de Córdoba* have held meetings in order to build an agenda for the institutionalisation of evaluation in the country.

### *Profession/Discipline*

The denomination and the understanding itself of evaluation as a profession are new in some countries (Rodríguez-Bilella et al., 2017) and strongly incipient in the context of Argentina. The practice of evaluating programmes and projects has been an activity adding to those activities usually carried out by academics, consultants and technicians. In this sense, the evaluation market in Argentina is limited and generally concentrated in the main cities of the country. Firms dedicated to evaluation often have also another main focus, such as popular education or social research. There are no scientific journals in the country with a focus on evaluation either, but there are others related to public administration, where there is a place for articles on evaluation.

This reality concerning a limited field for the discipline of evaluation is also to be seen in the absence of an arbitration board or association with the ability to regulate the performance of evaluators, both in terms of ethical issues, safety regulations and even salary issues. The *Red EvaluAR* (Evaluate Network)—Argentine Evaluation Network—briefly introduced in a previous section, is a Voluntary Organisation for Professional Evaluation that independently groups professionals and experts from Argentina with the aim of improving their practices and contributing to the local–global debate on the development of evaluation capacities. Its character is rather that of a professional growth network than one oriented towards disciplinary regulation. It does not rely on its own evaluation standards, but it has disseminated and highlighted the ‘*Estándares de Evaluación para América Latina y el Caribe*’ (Evaluation Standards for Latin America and the Caribbean) (Rodríguez-Bilella et al., 2016), published by ReLAC.

One of the actions of *Red EvaluAR* that proved significant in order to get to know the reality of evaluation in the country in greater detail was the research and publication of the ‘*Mapa Diagnóstico de la Evaluación en Argentina*’ (Diagnostic Map of Evaluation in Argentina, Aquilino & Amaya, 2015), which addressed both the reality of the national and sub-national system, taking as a case study the city of Buenos Aires and the

provinces of Salta, Mendoza, Corrientes and Córdoba. Aspects related to the understanding public sector actors have regarding the function and role of evaluation; the identification and analysis of national and provincial regulations in force within the scope of each province; the review of the design quality of plans and programmes to be evaluated; and the analysis regarding the perception of each province about the importance of strengthening evaluation capacities and its institutionalisation were highlighted through a qualitative strategy.

In conjunction with the calls to celebrate the Evaluation Week in Latin America and the Caribbean, based on an initiative of the *Centro CLEAR-LAC* (CLEAR-LAC Centre), the *Red EvaluAR* co-organised the first dialogue towards the institutionalisation of evaluation in Argentina in 2017, which aimed to nurture a broad and participatory debate about the relevance, need and potential of having public policy evaluation institutions. For this purpose, national, provincial and municipal public officials, academics, international organisations and civil society organisations were invited to discuss the institutionalisation of evaluation in countries with federal government systems and different evaluation experiences were contrasted.

### *Compliance with Quality Standards and Obligations*

The practice of evaluation in Argentina is not regulated as such by any professional entity. Evaluations carried out within the governmental framework must have the validation or consent of the instance having requested it, which is not usually guided by explicit norms or quality standards. Civil society organisations (NGOs, foundations) requesting evaluations often merely check on their concordance with the ToR included in the call for evaluation, as well as on issues related to its technical quality.

The document referring to the '*Estándares de Evaluación para América Latina y el Caribe*' (Evaluation standards for Latin America and the Caribbean, Rodríguez-Bilella et al., 2016) has been disseminated mainly among evaluators and has not impacted decision makers or clients. In this sense, a demand for standards to be followed is not to be seen among organisations, customers or evaluators.



## CONCLUSIONS

The institutionalisation of evaluation has historically been marginal within the framework of the priorities set by the different governments of Argentina. Together with the lack of a regulatory framework having an integral vision of evaluation in the national state—definition of priorities, obligations, responsibilities and evaluation standards—those norms or regulations that are close to monitoring and evaluation issues concentrate as a priority on realities focusing on control and monitoring of spending (accounting audit) rather than on learning and contributions to planning, thus becoming irrelevant to systematically challenge the development results achieved.

Although there are regulations providing technical bases for monitoring and evaluating government interventions, which mainly focus on the perspective of managing for results, performance evaluation and expense monitoring, legal and administrative regulation of evaluations and the use of their results reflect ambiguity and uncertainty concerning the importance of evaluation in the Argentinian Federal State. In this way, the different sectoral actions reflected in the monitoring and evaluation units existing in different national ministries constitute fragmented and disjointed experiences, lacking an integral vision giving them a strategic meaning and favouring and enhancing both promotion and use and capacity building in evaluation (Acuña et al., 2016).

The creation of evaluative structures has not been a strategically planned process within the framework of the national state in Argentina, which results in very different institutional capacities to design, plan, implement, monitor and evaluate policies among ministries, government agencies, provinces and municipalities (Aquilino et al., 2017). Consequently, monitoring and evaluation actions are subject to the political initiative and the technical will of people responsible for managing plans and programmes as well as to the demands of the funding agencies, setting up a scenario where very different experiences regarding quality and technical rigour coexist. Without a national policy aimed at ordering the monitoring and evaluation functions of state actions, evaluation as a policy depends more on the will of individual actors than on the institutional conviction of the state.

Evaluations of results usually do not circulate or are not disseminated beyond the scope of the administrative organisation that generated them (secretariat, management, institute), thus they remain mostly unknown to

other state agencies as well as to civil society. In relation to the latter, there is no interest regarding citizen participation in public policy evaluation processes to be seen, which would contribute to providing greater transparency to public management as well as strengthening accountability actions by those responsible for executing programmes and policies.

Initiatives such as those expressed by training programmes regarding evaluation and the consolidation of a professional network of evaluators provide inputs from civil society to move towards institutionalisation processes and integrality of the evaluation system in Argentina. The challenge is to overcome the character of initiatives isolated from each other and from state management, in order to generate articulations and synergies between these actors participating in the broad field of evaluation in Argentina.

### LIST OF ABBREVIATIONS

CONEAU	National Commission for University Evaluation and Accreditation
PRONEP	National Programme for Evaluation and Strengthening of Science, Technology and Innovation Programmes
ReLAC	Monitoring, Evaluation and Systematisation Network in Latin America and the Caribbean
SIEMPRO	System of Information, Evaluation and Monitoring of Social Programmes System of Information, Evaluation and Monitoring of Social Programmes
ToR	Terms of Reference

### REFERENCES

- Acuña, C. H., Martínez Nogueira, R., Rubio, J., & Potenza, F. (2016). La evaluación de políticas públicas en la Argentina: Sentido, actualidad y perspectivas. *Documento De Trabajo IIIEP*, 13, 1–37.
- Aquilino, N., & Amaya, P. (2015). *Mapa diagnóstico de la evaluación en la Argentina*. Red Argentina de Evaluación–EvaluAR. 1–72.
- Aquilino, N., Arias, E., Estévez, S., & Suaya, A. (2015). Hacia un análisis de evaluabilidad de planes y programas sociales: Un estudio sobre 33 iniciativas implementadas en Argentina. *Studia Politica*, 34, 37–72.

- Aquilino, N., Ballescá, M., Potenza, F., & Rubio, J. (2017). ¿Todos los caminos conducen a Roma? Análisis comparado en la institucionalización de la evaluación. *Documento De Trabajo CIPPEC*, 159, 1–31.
- Aquilino, N., Pomares, J., Suaya, A., & Page, M. (2016). *Déficits en la Rendición de Cuentas Horizontal en la Argentina: una Historia de dos Mundos*. CIPPEC.
- Canievsky, C. (2007). *Diagnóstico de los sistemas gubernamentales de monitoreo y evaluación en Argentina*. Banco Mundial; CLAD.
- Neirotti, N. (2000, October). *Reflexiones sobre la práctica de la evaluación de programas sociales en Argentina (1995–1999)*. Congreso Internacional sobre la Reforma del Estado y de la Administración Pública.
- Neirotti, N. (2001, November). *La función de evaluación de programas sociales en Chile, Brasil y Argentina*. Congreso Internacional del CLAD sobre la Reforma del Estado y de la Administración Pública.
- Neirotti, N. (2015). Hacia un nuevo paradigma en evaluación de políticas públicas. *La evaluación de las políticas públicas: Reflexiones y experiencias en el escenario actual de transformaciones del Estado. Remedios de Escalada: Editorial de la UNLa*, 1–14.
- Neirotti, N., Brissón, M. E., & Mattalini, M. (2015). Tiempo de retorno del Estado: realidades y desafíos del seguimiento y la evaluación de políticas y programas en Argentina. In G. Pérez Yahuarán, & C. Maldonado Trujillo (Eds.), *Panorama de los sistemas nacionales de seguimiento y evaluación de programas y políticas públicas* (pp. 31–70). CIDE.
- Ortale, S. (2015). Presentación de la Red de Evaluación Argentina (EvaluAR). *Studia Política*, 35, 173–176.
- Potenza, M. F. (2018). *Monitoreo y Evaluación de Programas de Inversión Social Privada*. Grupo de Fundaciones y Empresas.
- Rodríguez-Bilella, P., Valencia, S. M., Alvarez, L. S., Klier, S. D., Hernández, A. L. G., & Tapella, E. (2016). *Estándares de evaluación para América Latina y el Caribe*. Akian Grafica Editora.
- Rodríguez-Bilella, P. (2017). Significance of Voluntary Organizations for Professional Evaluation (VOPES) for the dissemination and professionalization of evaluation. *Zeitschrift Für Evaluation*, 16(2), 210–218.
- Stockmann, R., & Meyer, W. (2016). *Evaluación: una introducción teórico-metodológica*. Editorial UCR.
- Vinocur, P. A., & Halperin, L. (2004). Pobreza y políticas sociales en Argentina de los años noventa. *CEPAL*, 85, 1–84.
- Zaltsman, A. (2006). *Experience with institutionalizing monitoring and evaluation systems in five Latin American countries: Argentina, Chile, Colombia, Costa Rica and Uruguay*. Independent Evaluation Group ECD (Working Paper No. 16). World Bank.



## Evaluation in Bolivia

*Silvia Tatiana Salinas Mulder, Martha Lanza Meneses,  
Scarleth Flores Calle, María Dolores Castro Mantilla,  
and Paul Mauricio Villarroel Via*

### GENERAL OVERVIEW

Located at the heart of South America, the Plurinational State of Bolivia is constituted as a unitary, plurinational, free, independent, sovereign, democratic, intercultural and decentralised social state based on the rule of law, which acknowledges autonomous communities and regions within it. In addition, its government takes a participatory, representative and community democratic form, in equal terms and conditions for both men and women.

---

The authors of this chapter article belong to the REDMEBOL Coordinating Committee. Paul Villarroel is Associate Researcher of Aru Foundation.

---

S. T. Salinas Mulder (✉) · M. Lanza Meneses · S. F. Calle ·  
M. D. Castro Mantilla  
Red de Monitoreo y Evaluación de Bolivia (REDMEBOL), La Paz, Bolivia  
P. M. Villarroel Via  
Fundación ARU, La Paz, Bolivia

© The Author(s), under exclusive license to Springer Nature  
Switzerland AG 2022

65

R. Stockmann et al. (eds.), *The Institutionalisation of Evaluation in the Americas*,

[https://doi.org/10.1007/978-3-030-81139-6\\_3](https://doi.org/10.1007/978-3-030-81139-6_3)

This chapter is presented as a continuation of the work carried out by the Network for Monitoring and Evaluation of Bolivia (REDMEBOL) in relation to the institutionalisation of evaluation in the country. Among its milestones, the latter includes the publication of a diagnosis on the Institutionalisation of Evaluation in Bolivia (REDMEBOL, 2018).

Primary and secondary sources were used for the preparation of this document. It involved documentary review, seven interviews conducted and an online survey, which gathered a total of 29 responses from various state actors, as well as from civil society, the private sector, the academic community and independent professionals (consultants).

## INSTITUTIONAL STRUCTURES AND PROCESSES (POLITICAL SYSTEM)

The institutionalisation of evaluation in Bolivia is nowadays at an early stage. However, important progress has been made in recent years, both in the public sector and in some civil society professional groups, which show the beginning of a path towards the goal.

The understanding of the current situation of institutionalisation in Bolivia must be based on the analysis of the existing regulations in the country, as well as the practice and effective use of evaluations throughout the public policy cycle.

### *Evaluation Regulations*

The first dimension of analysis focuses on those regulations in force in the country, which govern public sector activities linked to evaluation obligations and planning systems.

The Executive Branch of the Bolivian State is regulated by Supreme Decree No. 29894 issued on February 7th (Estado Plurinacional de Bolivia [2009b]). This regulation establishes the structure of the executive branch and defines the main powers and obligations of each authority composing it. In this sense, the decree establishes responsibilities regarding public policy evaluation activities for each of its components.

Title I of General Provisions defines in its third chapter the main tasks assigned to the ministers and public servants of the Plurinational State of Bolivia. Article 14, subsection 9, establishes the following attributions and obligations:

To promote and implement, in coordination with social movements, public policies and activities concerning the evaluation and control of public management

It can be seen that the responsibility for evaluation and control regarding public policies lies primarily with the highest executive authorities of each of the ministries belonging to the executive organ of the State.

In Bolivia, the Ministry of the Presidency (MDP) is the coordinating body between the president of the Plurinational State and the different ministries. Within the same decree, Article 23 establishes that the Vice Ministry of Territorial Government Coordination and Management, under this ministry, has the authority to “carry out monitoring and evaluation activities and analysis concerning strategic projects within the different ministries and their sustainable impact”. In the same way, many other ministries have the authority to carry out monitoring and evaluation activities related to the policies and projects they carry out.

Although Bolivian ministries have powers to carry out monitoring and evaluation activities concerning sectoral policies, the Ministry of Development Planning concentrates the main obligation in this regard.

Supreme Decree No. 29894 establishes monitoring and evaluation activities regarding the Economic and Social Development Plan (PDES), as one of the obligations of the highest executive authority of this ministry, in coordination with other State institutions. On the other hand, the authority has the obligation to carry out monitoring and evaluation activities concerning the Comprehensive Planning System of the Plurinational State (SPIE), in addition to monitoring and evaluating the State Investment and Financing System.

One of the most important powers of the MPD, related to the planning, monitoring and evaluation of public policies, is the role of full guardian of the Economic Policy Analysis Unit (UDAPE). According to the organisational manual (UDAPE, 2014), this decentralised institution is responsible for the analysis of economic and social policies, the sub-directorates belonging to this unit have the obligation to monitor and evaluate macroeconomic and social policies.

Regulations analysed so far focus on the general powers held by those bearing principal responsibility for monitoring and evaluation. However, Supreme Decree No. 29894 only regulates functions within the executive body, but it does not detail specific regulations on monitoring and evaluation activities within the state.

On the other hand, after carrying out an analysis of the specific regulation of UDAPE, as a decentralised institution, it is not possible to observe detailed regulations on the production and/or use of evaluation within the different sectors.

On 21 January 2016, within the framework of national regulations concerning the planning, monitoring and evaluation of public policies, Law No. 777 was enacted, aiming to establish the SPIE (Estado Plurinacional de Bolivia, 2016). As mentioned earlier, the MPD is the governing body responsible for monitoring and evaluating the system.

According to Article 2, the SPIE aims to articulate norms, subsystems, processes, methodologies, mechanisms and procedures regarding integral state planning in the short, medium and long terms. With an autonomous vision of sovereignty and dignity, this instrument redirects and integrates the processes of development planning and evaluation, at all state levels and sectors; as it articulates sectoral and territorial planning processes and integrates long-, medium- and short-term planning, within the framework of ‘*Vivir Bien*’ (engl.: *Living Well*).

The subsystems that make up the SPIE and articulate the different guidelines are the following:

- Planning
- Public Investment and External Financing for Integral Development
- Comprehensive Monitoring and Evaluation of Plans.

We highlight some important articles of Law No. 777:

According to article 3, one of the purposes of the SPIE is: “To carry out a comprehensive monitoring and evaluation of planning, based on goals, results and actions, providing timely information for public management decisions”.

Article 11 establishes the Cyclical process of integral state planning targeting *Vivir Bien*. This process of permanent qualitative progress begins with the formulation of plans within the different State sectors. Once these are approved, the allocation of resources concerning integral policy, programme and/or project implementation is carried out. Within this process, regulations establish follow-up as a key element for subsequent evaluation and adjustment. It is important to note that there is a line of continuous communication and feedback throughout the process,

fulfilling the cycle once evaluations and adjustments are considered for the formulation of plans in order to start the process again.

Article 7 refers to the obligation all State institutions making part of the system have to “[...] carry out control, monitoring and evaluation to achieve the goals, results and actions contained in their plans [...]”. This shows that there is a clear intention to generate evaluation processes as planning tools concerning public policies.

The Integral Plan Monitoring and Evaluation Subsystem is established in Chapter 3, which defines it as “[...] the set of guidelines, methodologies, procedures and technical instruments aimed at systematising, analysing and evaluating compliance of goals, results and actions of long, medium- and short-term plans [...]”.

In reference SPIE evaluation, article 31 establishes that the integral evaluation of plans will be carried out within the framework of a quantitative and/or qualitative midterm assessment. On the other hand, it defines the entities holding major responsibility for the integral evaluation of all plans making up the SPIE. Regulations also establish the periodicity and types of evaluations required, regarding evaluations of actions, results and impacts as the most important.

Finally, it is necessary to mention that the SPIE articulates different plans—each one holding specific evaluation protocols—which altogether organise the policies and/or specific strategies of the Plurinational State of Bolivia in the long, medium and short terms.

### *Sector Regulations*

The Integral Development Plans for *Vivir Bien* (PSDI), which are part of the SPIE, govern the planning, monitoring and evaluation processes concerning the different sectors throughout the ‘Methodological Guidelines for the Formulation of Integral Development Sector Plans for Vivir Bien’ (span.: *Lineamientos metodológicos para la formulación de planes sectoriales de desarrollo integral para ‘Vivir Bien’ PSDI*) (Ministerio de Planificación del Desarrollo, 2016). These instruments, which are operative in nature, must articulate the planning of each sector under the guidelines of the Institutional Strategic Plan and Public Business Plans.

For each PSDI, a diagnosis must be made based on comparative evaluations of the sector in recent years and evaluations its current status. All sectors must establish pillars, goals, results and actions in order to allow the monitoring of policies. In addition, this regulation establishes that the monitoring of the goals, results and actions of the plan will be carried out



on an annual basis and its impact evaluation will be conducted at midterm and at the end of the five-year period.

In addition to the PSDIs defined in each sector, we mention some examples of sector regulations related to evaluation processes. For this purpose, we considered and analysed the education and health sectors, because they made considerable progress in terms of planning, monitoring and evaluation of their programmes.

### *Education*

On 20 December 2010, Education Law No. 070 '*Avelino Siñani - Elizardo Pérez*' was enacted (Asamblea Legislativa Plurinacional, 2010). Section V of this law includes the norms established for university evaluation and accreditation and determines the powers of the Plurinational Agency for the Evaluation and Accreditation of Higher University Education.

Article 83 of the law establishes the creation of the Plurinational Observatory of Educational Quality, which is responsible for monitoring and evaluating the education system in the Regular, Alternative and Special subsystems. Likewise, Article 91 regulates community social participation and determines its role in the planning, control, monitoring and evaluation of the educational process.

### *Health*

According to the organisational manual of the Ministry of Health approved by Ministerial Resolution No. 0092 (Ministerio de Salud, 2012) the Planning Department is responsible for coordinating and evaluating the main goals and objectives of the health sector. Additionally, it has the authority to coordinate and evaluate the results of the National Health Information System. This system allows the collection of information from the health sector by generating statistical health information for each management period, and thus feed the monitoring, evaluation and decision-making processes. The statistical information generated in this system should allow the conduction of regular evaluations of the health situation in Bolivia, through the joint work of the Ministry of Health and other State institutions.

### *Regulations and Characteristics of Evaluation*

The analysis of the regulations in force in Bolivia allows us to observe that norms have an instructive approach regarding the responsibilities of

the main institutions and officials of the Plurinational State of Bolivia. Since 2016, Law No. 777 and the SPIE explain in greater detail the characteristics that these monitoring and evaluation processes must have.

Regulations focus especially on the periodicity of evaluations and the tools that are necessary to conduct them, however, they do not address the required quality, specific methodologies or the independence that evaluations carried out should have.

### *Evaluation in Parliamentary Structures*

Within the parliamentary structure of the Plurinational State of Bolivia, called the Plurinational Legislative Assembly, regulations established for both the Senate Chamber and the Chamber of Deputies can be identified. However, specific guidelines on the use of evaluation or the need to implement monitoring processes for public policies cannot be evidenced.

The two Chambers constitute the supervisory entities of the State organs. The General Regulations of the Senate (Estado Plurinacional de Bolivia, 2012a), which regulate Articles 158 and 160 of the Political Constitution of the State (2009a), establish that Senators have the authority to supervise the different activities of the State and request reports to exercise this responsibility. However, there is no regulation specifying the nature of these reports or their direct relationship with public policy evaluation processes.

In the same way, the General Regulations of the Chamber of Deputies (Estado Plurinacional de Bolivia, 2012b), define their audit capacity and establish the procedures required to exercise this right. As in the case of the Senate, there is no regulation specifying the use of evaluations or similar tools to carry out its work.

### *Evaluation Practice*

At this point, we will analyse the relationship between what is established by Bolivian regulations and the practice and effective use of evaluation.

Through the implementation of the Operations Programming System, within the framework established in Article 27 of Law No. 1178, all state entities prepare their Annual Operational Plan (POA) and regulate their monitoring processes and evaluation (El Honorable Congreso Nacional, 1990). Non-compliance can generate responsibilities for the public service and sanctions. In this context, it can be pointed out that public entities carry out evaluations based on the POA, with emphasis on effectiveness

and efficiency in achieving planned results (mainly oriented to physical-budgetary execution at the product/outcome and process level).

Currently, within the framework of Law No. 777, the MPD is responsible for planning and monitoring. For this purpose, it verifies that Public Sector Planning is aligned with the PDES 2016–2020, and monitors compliance with results and goals. The National Public Investment System depends on the MPD, which determines the procedures and responsibilities for public investment projects.

The National Statistics Institute (INE) generates information for integral planning at the different state levels and sectors. It is currently adjusting its methodology for the change of base year.

For the first time, the MPD has carried out a midterm evaluation of the fulfilment of the goals of the PDES 2016–2020. Development plans at sector, institutional and territorial level (departmental, municipal, regional and indigenous peasant native) were evaluated.

As part of the implementation of the SPIE, the INFO-SPIE has been created in the perspective of providing a single source of official information accessible to the public through its website,<sup>1</sup> whereby the Autonomous Territorial Entities develop their planning, monitoring and evaluation processes. This situation tends to improve processes.

The UDAPE, which depends on the MPD, carries out impact evaluations of social programmes, analysing the effects attained regarding the improvement of well-being. UDAPE has evaluated national policies such as Renta Dignidad, Juana Azurduy Bonus, Juancito Pinto Bonus and Empleo Digno (Decent Employment), among others. Evaluations are mainly carried out funded by the General Treasury of the Nation and External Credit, with external donations representing around 10%. This unit was responsible for preparing the Millennium Development Goals reports and currently focuses on the 2025 Patriotic Agenda and its relationship with the Sustainable Development Goals.

Generally, funding agencies carry out midterm, final and in some cases impact evaluations, in coordination with the corresponding sector ministries. For this purpose, they include the corresponding budget in programmes and projects.

The evaluations of national programmes are midterm and final evaluations, the tendency is to evaluate at product/result and process level.

<sup>1</sup> The website can be found here: <http://si-spie.planificacion.gob.bo/>.

In the case of sectoral evaluation, some ministries have institutionalised monitoring and evaluation bodies responsible for evaluations and reporting indicators; such as the Plurinational Agro-Environmental Observatory under the Ministry of Rural Development and Land (MDRyT), the Productive Analysis Unit under the Ministry of Productive Development and Plural Economics (MDPyEP), the Plurinational Observatory for Education Quality under the Ministry of Education. These instances have been institutionalised and operate with their own resources after they were created with the support of cooperation.

It is important to highlight the planning in multisectoral programmes, in which several State ministries participate. These, in turn, develop monitoring processes for indicators and periodic evaluations.

In 2015, the Electronic Government Agency for Information and Communication Technologies was created to strengthen monitoring systems permanently. Access to information has been improved through the General Personal Identification System (SEGIP) and the Civic Registration Service, and the Ministry of Health is working to have vital statistics in coordination with SEGIP.

The Central Bank of Bolivia and the Financial System Supervision Authority have developed solid and structured monitoring and evaluation processes in the financial area, information generated is the basis for the creation of public policy.

On the other hand, the Office of the Comptroller General of the State performs the Internal Control of all public entities through audits with emphasis on efficiency and reliability, mainly in budgetary execution and the performance rate of public servants.

### *Use of Evaluations*

Regarding the use of evaluations, although it cannot be stated categorically that these are ‘good performance’ sectors with respect to the use of evaluation results, we mention some cases that exemplify these processes.

In the case of the health sector, it carries out evaluations of public policies with an integral and qualitative-quantitative character, analysing the effects of these up to departmental levels. It should be noted that planning and monitoring processes in the health sector provide instances for citizens participation through the Local Health Councils, Municipal Health Councils and Departmental Health Councils. The planning processes are carried out according to social determinants, which allows to

visualise the results of health policies in an integral way, linking different sectors.

Specifically, the Multisectoral Zero Malnutrition Programme, which aims to eradicate malnutrition in children under five, has been evaluated. This programme is part of the Social Protection and Community Integral Development Policy and was implemented by the National Food and Nutrition Council in which six ministries participate. The results of the achievements-based evaluation have allowed us to rethink the policy by applying the same modality but extending its reach to the entire population, including the elderly. The evaluation was carried out based on information from the Demography and Health Survey 2016.

The MDPyEP has developed an internal evaluation of those activities linked to dairy production and processing (*complejo lácteo* in Spanish) according to the incidence of the actions coordinated with the MDRyT. The results/effects of the evaluation include, among others, the improvement of credit access for dairy producers, mostly credits from the Productive Development Bank, better distribution of income in the dairy sector after the creation of the *ProLeche* (Pro Dairy) Fund to subsidise the price of milk, and the strengthening of production in order to respond to the policy of increasing milk consumption per capita (the national goal is two glasses of milk per day). Based on these results, the sector policy has been adjusted to improve the processing technology of dairy products, generating greater added value and improving market access for producers.

UDAPE has carried out the evaluation of the Juana Azurduy Programme, aimed at reducing maternal and neonatal mortality. As a result of the recommendations of the aforementioned study, budgetary adjustments were made to reduce programme costs. Regarding the use of evaluation results arising from this type of national programmes, the impact goal established by the UDAPE regulatory framework is to redefine the policies' objectives, goals and results and even to reorient the sector. As for the quality of evaluation, UDAPE uses international rigour standards in the evaluations.

As mentioned in the introduction, in order to determine the use of evaluation, REDMEBOL implemented an online survey aimed at people who develop their activities in private companies, the academic community, independent people and public servants. 3.45% of the 29 people who participated responded that they 'never' used evaluation, 31.03% that they 'rarely' did, 37.93% that they 'often' did and 27.59% said they 'always'

used it. It is worth mentioning that 72.41% of the people who answered the questionnaire represented the private sector and civil society.

Regarding evaluation use in its professional activity, 65.5% of respondents affirm that they evaluate 'often' or 'always' (people who carry out their activity in state entities, academia, NGOs and independents); 50% say they use evaluations to reorient projects or as a tool for learning good practices and determining lessons learned; and 16% say they use them because it is a requirement of the funder.

82% of respondents agree that those who use evaluation results are the managers of programmes or projects, only 3.4% believe that political leaders use evaluation results (perception of civil society / NGOs).

In general, those who use evaluations argue that they do so to make informed, basic decisions for project management and financing; redirect projects, make adjustments and/or define project continuity. They also use evaluation as a learning tool for technical teams, funders, population involved, authorities, etc. It is also used to develop future plans; monitoring of the activities carried out in the institution; timely decision making and to identify good practices and lessons learned in order to design future programmes and projects.

On the other hand, they consider that those who use evaluations usually do so due to external impositions of the project as a requirement of funders, since this can even be a condition to move on towards another stage of work.

Project managers use evaluations to refer to data that supports their actions or schedules, provided they are properly trained technicians with basic evaluation skills and knowledge about its usefulness and/or are linked to the design or implementation of programmes and projects.

Those who consider that political leaders use evaluations believe that they do it for their own benefit, depending on whether evaluations support a political situation or not, but they never use them to make technical decisions. Every political situation needs an impact speech, so they use data to nurture speeches of state interpellation or to support it.

Regarding the limitations in the use of evaluations, 33% of respondents consider that they are due both to the lack of custom and public indifference; 28% think they can be explained because of the lack of legitimacy of evaluations and 11% believe that there is no access to evaluations.

In general, it can be noted that the MDP, the UDAPE, the INE and sectoral state entities use evaluations to formulate or adjust national policies, while policy operators use these to improve their management.

Advances in both regulations and evaluation practice are important; however, challenges are also substantive, especially considering weaknesses related to (i) a culture that identifies evaluation as a control mechanism rather than as a tool for transformation; (ii) the weak evaluation training offer for human resources; finally; and (iii) the high mobility of personnel in the public administration, among other aspects. Perhaps the use and quality of evaluations present greater challenges to the extent that these aspects have not been considered in current regulations, as indicated in section ‘Evaluation Regulations’, and in general depend on personal decisions in the case of professionals, and political decisions regarding public management. Finally, it will be necessary to consider that the SPIE was just approved in 2016 and its proper application with respect to the evaluation processes included in these regulations may depend to a large extent on the challenges already mentioned.

## DISSEMINATION IN SOCIETY AND ACCEPTANCE (SOCIAL SYSTEM)

### *Institutionalised Use of Evaluations by Civil Society*

In general, there is a negative perception or lack of knowledge about the use of evaluations for referendums or other political processes and decisions, at the national, departmental and/or sub-national level. Although a percentage greater than 25% of the survey responses refer to the eventual use of evaluations for referendums and other decisions of a political and community-based nature, there is no qualitative information to explain the percentage of responses classified under ‘sometimes’.

The associated problems or obstacles have to do first with the availability of evaluations, secondly with their dissemination and accessibility and, thirdly, with the absence of a culture promoting both governmental and civil society actions, and even from the private-business sector, based on evaluations and, more generally, on evidence: “In general, in Bolivia there is a very weak culture of debate based on information; there is no habit of consumption and use of the data provided by evaluations” (online survey respondent).

In general, evaluations are considered ‘made to be fulfilled’, have little diffusion in terms of their results and are far from being considered an investment and a useful tool for learning and decision making.

There is a gap between actions which are considered political and the more academic or evidence field. Political actions are not perceived as a response to certain evidence, and evaluation is also not perceived as a political action, in the sense of its transformative potential: “It seems to me that referendums and other processes reach civil society highly mediated and contaminated with political pressure and they do not have much to do with evaluations” (online survey respondent).

Limitations regarding the use of data, which have to do with culture, but also with lack of knowledge and weak skills ‘to carry out and interpret evaluations’, inevitably refer to a low training offer in the field of evaluation in the country.

When it is not a question of final evaluations, evaluations are not always accompanied by flexible budgets, nor of a real will to make substantive changes, so recommendations frequently remain as such. On the other hand, it is not possible to evaluate the implementation of regulatory frameworks when laws lack a budget. It is worth mentioning there are evaluations of the implementation of norms conducted by civil society, especially concerning the provision of public services.

Regarding the role of civil society, its participation is mostly functional, as an information provider, that is, as a set of respondents (see Table 3.1):

**Table 3.1** What role does civil society normally perform in evaluations?<sup>2</sup> (Own development)

<i>Answer</i>	<i>Frequency</i>	<i>Percentage</i>
Customers	5	17
Respondents	25	86
Users of evaluation results	6	21
People providing input and concepts based upon their life experience	1	3

<sup>2</sup> In this question, it was possible to choose several options.

It is not possible to estimate the number of evaluations concerning NGOs in Bolivia, or other instances of civil society, based on these roles. This information is not centralised, and in general it is not publicly accessible.



Regarding the role of ‘respondents’, mostly assigned to civil society, it is also worth questioning the quality and value of this participation. The latter has to do with the questions asked, with the assumptions made regarding people’s poor knowledge, and with the consequent undervaluation of their input, which is usually not even considered. Thus, participation is confined to a formal process, in order to fulfil a requirement.

Just as wide dissemination processes regarding the few evaluations carried out are limited, those of feedback to people that participate and contribute in evaluations are practically non-existent, and so is also their impact on the decisions that have to do with the dissemination and uses of the evaluation, therefore. This has to do also with report formats and inaccessible language: “There are no understandable versions for people. People’s language is not used [...] Even if you distribute a photocopy to each family it will not reach them” (interview respondent).

*Perception and Public Discussion of Evaluation and Evaluation Results*

The majority perception is that there is no culture of evaluation in Bolivia, a fact that reconfirms the initial findings of the diagnosis on institutionalisation in Bolivia, undertaken by REDMEBOL in 2017 (see Table 3.2).

**Table 3.2** Do you think evaluation culture does exist in the country? (Own development)

<i>Answer</i>	<i>Frequency</i>	<i>Percentage</i>
Yes	0	0#
Partially	11	37
No	18	62
Total	29	100

The existence of a culture means that the topic is positioned or taken over by society, and that there is no need for external factors or actors to determine its validity. In Bolivia, the issue of evaluation is mostly maintained as a requirement of international cooperation, which means that “evaluation is conducted to meet a requirement, but not necessarily because you are thinking about making changes in your institutional practice” (interview respondent). This is associated with little appreciation for

data, a weak orientation towards learning and acceptance of criticism, as well as towards a culture of accountability, transparency, and focus on results. It also refers to the “absence of an institutional and citizen culture of anticipation and corrections of procedure” (online survey respondent).

As mentioned earlier, the limited availability and access to the few existing evaluations is a problem at the level of the State in general, which has to do largely with politicisation and ‘*secret evaluations*’, emphasising, however, that there is also no culture in private organisations—encompassing civil society—to socialise and enable permanent access to evaluations (e.g. on their web pages) (see Table 3.3).

**Table 3.3** Are evaluation reports (full version) available to the public? (Own development)

<i>Public sector</i>	<i>Frequency</i>	<i>Percentage</i>	<i>Private sector</i>	<i>Frequency</i>	<i>Percentage</i>
Yes	2	7	Yes	0	0
Sometimes	10	35	Sometimes	19	65
No	14	48	No	8	28
I don’t know	3	10	I don’t know	2	7
Total	29	100	Total	29	100

Related to all of the above, the public discussion about the uses, benefits, quality and professionalisation of evaluation is virtually zero: “Since they do not have a culture of accountability and transparency of information, the organisations themselves make no effort to place evaluation results on the public agenda” (online survey respondent).

The latter occurs despite the fact that there is a percentage of responses categorised under ‘sometimes’, which amounts to 35% in the public sector. However, that percentage does not hide the eventual, cyclical and without a doubt non-institutionalised nature of the possibilities civil society has to access public evaluation reports. In the case of the public sector, the answers ‘sometimes’ practically double those of the private sector, showing greater availability and access in general, but also its eventual and temporary nature. This could also underline a general overview, which includes certain organisations that make their evaluations available as an institutional policy, while there is a majority group that does not do so on a regular basis.

Evaluation is not an issue on the agenda, “evaluation results are not considered as public value or something worth discussing” (online survey respondent) (see Table 3.4).

**Table 3.4** Are the results of current evaluations publicly discussed (Own development)

<i>Answer</i>	<i>Frequency</i>	<i>Percentage</i>
Always	0	0
Sometimes	12	41
Never	10	35
I don't know	7	24
Total	29	100

In general, an evaluation is not discussed, unless the subject it addresses is spectacular (capable of achieving spectacle). “The media are not interested in causes, they are rather interested in what is entertaining or consumed” (online survey respondent). The latter would also depend on the results: “In my experience, if they are positive, that is, if all hypotheses worked out perfectly (something rare in fact), the chances that it will happen increase” (online survey respondent).

Regarding mass media, they are not very familiar with evaluation, so they do not promote or debate it publicly. Some of them broadcast debate and discussion programmes, where the social evaluation of certain problems is made visible.

It is important to mention REDMEBOL in this section. It is for certain the best-known actor carrying out various activities related to evaluation. This network was born in 2004 thanks to an inter-institutional agreement signed by four institutions: *Enlace Consultores en Desarrollo*; Save the Children; the Centre for Studies and Projects (CEP) and the Bolivian Centre for Multidisciplinary Studies (CEBEM).

Since 2016, a process of impulse, strengthening and repositioning of the network has begun at the national, regional and global levels. This includes a strategic planning process that implies important definitions such as the mission and vision of REDMEBOL, which emphasise the diverse and multi-actor composition of the network and its commitment to evaluation as a tool for social transformation. This framework stresses the promotion of ethics, innovation, the qualification and institutionalisation of processes and the use of monitoring and evaluation.

Within that framework, different initiatives and activities have been carried out aiming to promote a culture of evaluation, starting with the realisation of a diagnosis on institutionalisation, as well as a joint project with the Argentine Evaluation Network (EvaluAR), aimed at promoting institutionalisation. Likewise, following an approach focused on coordination and collaboration, alliances with national public and private institutions, international cooperation and independent evaluators have been developed and strengthened. At international level, these have also been built with other networks such as the Monitoring, Evaluation and Systematisation Network of Latin America and the Caribbean (ReLAC), the Network of Latin American and Caribbean Women in Organisation Management RedWim, the Argentine Evaluation Network (EvaluAr) and the International Public Policy Evaluation Network (RIEPP).

Although REDMEBOL is based in La Paz, after this experience other sub-national affiliated networks have been created, though their action and structure are independent. These networks are found in other departments such as: Santa Cruz, Oruro and—recently—Potosí. Sub-national spaces undoubtedly merit strengthening, but it should be noted that it is one of the few—if not the only—national network in the region that has promoted sub-national chapters.

As part of this process of inclusion and work towards institutionalisation, in 2017 REDMEBOL also promoted the creation of the Team of Young Evaluators of Bolivia—which is actually part of EvalYouth. This group is both interested in evaluation training and committed to the dissemination and contribution to the building of an evaluation culture in the country.

### *The Demand for Evaluation in Civil Society*

Demand is an action that emanates from a (perception of) need, importance, search for justice, etc. From previous sections, it can be derived that demand for evaluation is weak, in a context where evaluation culture is absent and expertise is not recognised as a public value (see Table 3.5).

Possible causes for this are a weak socialisation and implementation of the regulatory framework, orientation to results and informative transparency; also, interests and political struggle, and the incipient role of the media, since very often social media are now the channel through which civil society demands evaluation results. In an ‘institutionalised’ way, evaluations are associated with accountability events (economic, of results) at

**Table 3.5** According to your experience and knowledge, do individual citizens, civil society organisations, private companies or other actors require evaluations of, for example, political leaders? (Own development)

<i>Answer</i>	<i>Frequency</i>	<i>Percentage</i>
Always	1	4
Sometimes	14	48
Never	9	31
I don't know	5	17
Total	29	100

the end of negotiations. As an interesting related note, which could be the subject of a research project, “peasant communities and indigenous peoples have their ways of evaluating their leaders and their work” (online survey respondent).

PROFESSIONALISATION (PROFESSIONALISATION SYSTEM)

*Academic Courses, Training, Others*

There are no formal academic offers aimed at professionalising monitoring and evaluation (M&E) in the country, that is, there is no academic training programme granting a higher diploma in the evaluation field. Rather, there are specialised M&E courses within other disciplines, as can be seen in the documentary review of postgraduate educational personal data of the most important universities in the country (REDMEBOL, 2017). In this review, the disciplines in which M&E issues are most included are: Rural Development, Economics, Business Administration, Engineering, Education and Health.

According to a review of their websites, the main universities offering masters’ degrees, diplomas and specialties related to monitoring and evaluation in the country are *Universidad Mayor de San Andrés* (La Paz), *Universidad Mayor de San Simón* (Cochabamba) and *Universidad René Gabriel Moreno* (Santa Cruz). Training offers are linked to environmental impact evaluation, development projects, higher education, among others. The main recipients of these offers are professionals from technical areas related to the programme with durations ranging from four months to two years.

Another link between the academic community and evaluation has been ‘curricular evaluation’. Since the beginning of 2000, Bolivian universities have included curricular evaluation as part of their internal policies in each university.

As a result of curricular evaluation, teachers–evaluators have been trained as part of evaluation teams, following international standards and parameters established by UNESCO (United Nations Educational, Scientific and Cultural Organisation). That is why external expertise and accreditation processes concerning faculties and careers in several public universities in Bolivia have been carried out (*Comité Ejecutivo de la Universidad Boliviana* (CEUB, 2003)). However, L. Oporto states that Bolivia has departed from international evaluations on educational quality in recent years, and no such measurements are conducted in this regard in the country. “While there are no quality indicators and standards, duly validated, you simply cannot evaluate the quality of the teaching programs” (Diario Página Siete, 2015).

In the virtual survey carried out for this publication, a lack of knowledge about the existence of university programmes of higher education for evaluators is shown, as can be seen in the following Table 3.6.

**Table 3.6** Are there university programmes of higher education for evaluators (Diplomas, Masters) in Bolivia? (Own development)

<i>Answer</i>	<i>Frequency</i>	<i>Percentage</i>
Yes	9	31
No	7	24
I don’t know	13	44
Total	29	100

In interviews with two academics and evaluation consultants, they mentioned that “in the non-academic field there are other forms of training that interested consultants or professionals can access, through institutions or organisations inside or outside the country. These trainings are usually oriented towards the staff of institutions. Independent consultants or professionals have to look for opportunities via the Internet” (interview respondent).

The interviewees mention that the majority of projects in which consultants work assume that the latter should already have been trained in evaluation and very few projects consider including the development of skills or competencies in this area.

In the virtual survey, the question about other academic or non-academic forms of evaluation training produced the following results (see Table 3.7):

**Table 3.7** Are there other forms of academic or non-academic training on evaluation? (e.g. e-learning, training of consultants, etc.)? (Own development)

<i>Answer</i>	<i>Frequency</i>	<i>Percentage</i>
Yes	16	55
No	2	7
I don't know	11	38
Total	29	100

Other training forms identified correspond to:

- E-learning
- Inter-American Development Bank (IDB)-Platform
- Webinars, online courses, conferences
- REDMEBOL
- EvalPartners
- NGOs

In 2017, during a binational workshop held to know the results of the diagnoses on the institutionalisation of evaluation in Bolivia and Argentina, carried out by REDMEBOL and EvaluAR, the group that addressed the development of individual and professional capacities in evaluation evidenced that evaluation does not occupy a prominent place in Latin America and the Caribbean yet, which contrasts with a growing demand for evaluations. The results of the diagnosis show the need to review, enhance and innovate approaches and evaluation methodologies with flexible and at the same time rigorous contents, mainly containing key transversal axes for gender equality, the exercise of rights, as well as ethical and culturally relevant positions.

Among the main conclusions presented by the working group, the fact that social and economic sciences are the careers most linked to evaluation stands out. There is a very limited education offer, training is limited to a

large number of short-term proposals such as diplomas, courses and workshops offered by international evaluation networks or other international organisations.

### *Profession/discipline*

The virtual survey showed that most respondents do not know or are not familiar with the existence of newsletters or other means of information that disseminate news, events, activities and others related to evaluation and monitoring. However, some research institutes of university postgraduate or undergraduate studies centres sporadically publish the results of policy or project evaluation.

It is worth mentioning that a good part of the survey participants referred to REDMEBOL as the most important means for these purposes. They also mentioned the RIEPP, the ReLAC and the Latin American and Caribbean Evaluation and Monitoring Network among the most important. Through REDMEBOL you can access different kinds of information about events, courses, scholarships, activities in the country and abroad, information widely disseminated through the Facebook page.<sup>3</sup> REDMEBOL also has three newsletters in physical and electronic formats, with the main objective of disseminating institutional activities and promoting evaluation and monitoring in the country. On the other hand, in recent years the RIEPP newsletter has represented a space through which news are disseminated from Bolivia and from which Bolivian evaluators also learn about activities and events in Latin America. Through its email, REDMEBOL is the main actor in charge of reaching more than 100 people linked to the network and disseminating the information through a list of emails.

The country does not have an institutionalised system of norms or principles to guide the actions of evaluators, usually these depend on the recruiting institutions for evaluations and/or the criteria thereof. As can be seen in the following Table 3.8, most of the participants who completed the online survey do not know if there are norms, principles or guides:

Although they mention that public accountability is regulated, the resting spaces are managed according to traditions and customs. As

<sup>3</sup> The website can be found here: <https://es-la.facebook.com/Redmebol/>.



**Table 3.8** Are there norms, guiding principles or others similar for evaluators in your country? Any guidelines developed by REDMEBOL that you know? (Own development)

<i>Answer</i>	<i>Frequency</i>	<i>Percentage</i>
Yes	3	10
No	3	10
I don't know	23	79
Total	29	100

mentioned above, REDMEBOL has proposed a series of values for its actions. It also raises and promotes an inclusive and transformative approach to evaluations, with a gender, intercultural and intergenerational approach.

According to the interviews and the survey, the majority of respondents state that the evaluation market is covered by independent evaluators and some recognised ‘Think Tanks’, such as the ARU Foundation and the Institute for Advanced Development Studies that stand out in the impact evaluation market. CEBEM stands out as an organisation specialised in acting as an intermediary and calling for independent evaluations of private institutions and international organisations. In turn, this institution offers sporadically courses for specialisation in evaluation and impact (see Table 3.9).

**Table 3.9** Would you say that the evaluation market in your country is mostly dominated by independent, consulting or scientific research institutes? (Own development)

<i>Type of organisation</i>	<i>Frequency</i>	<i>Percentage</i>
Independent		
1	16	55
2	9	31
3	4	13
Consulting firms		
1	9	31
2	13	44
3	7	24
Institutos de investigación		
1	6	20
2	9	31
3	14	48

For participants, research institutes are not relevant, and consulting firms are more oriented towards information gathering (diagnostics, studies) than for evaluation. Since evaluation is not institutionalised in the State, cooperation programmes and projects resort mostly to individual consultancies or companies that carry out this work.

There is no certification system for evaluators in the country. Nor is there an authority that can be requested to reconcile in case of dispute.

### *Compliance with Quality Standards and Obligations*

Compliance with quality standards and obligations responds to the internal mechanisms of the recruiting institutions, both public and private. The primary, secondary and higher education system has its regulations regarding the quality of education.

According to the CEUB, in relation to higher education, there is the General Regulation for the Evaluation and Accreditation of Careers and/or Programmes, which defines evaluation as “the process of collecting information that allows, once analysed and interpreted in the light of the referential framework, to issue value judgments on the operating conditions of the Career and/or Programme, as well as to account for its quality and relevance, leading to decision-making related to accreditation” (CEUB, 2003, Capítulo II, Artículo 3, pp. 583–584). The regulation aims to regulate the procedures and activities of the evaluation and accreditation processes of the faculties and/or programmes of the Bolivian University system.

A demand for evaluation standards made by contracting institutions towards the evaluators was not noted, despite there being a concern about this issue. International cooperation, including agencies of the United Nations System, follows its own regulations or evaluation standards.

## CONCLUSIONS

Bolivia presents important advances in the institutionalisation of evaluation, it has clearly defined governmental entities responsible for the planning, monitoring and evaluation processes of public policies. The country has developed specific regulations for planning, monitoring and comprehensive evaluation of short, medium and long terms, articulating the processes of sectoral and territorial planning based on the PDES 2016–2020, which is currently in midterm evaluation process. However,

challenges are also substantive, especially considering weaknesses related to (i) a culture that identifies evaluation as a control mechanism rather than as a tool for transformation; (ii) the weak offer of human resources training in evaluation; (iii) the limited use of evaluation results; and finally, (iv) the high staff turnover within the public administration. Perhaps the use and quality of evaluations present greater challenges to the extent that these aspects have not been considered in current regulations, and in general depend on personal decisions in the case of professionals, and on political decisions regarding public management. Finally, considering that Law No. 777 of the SPIE (State Integral Planning System) was approved in 2016, its proper application with respect to evaluation processes may depend to a large extent on the challenges already mentioned.

On the other hand, civil society in Bolivia generally has a negative perception or lack of knowledge about access to evaluations and, consequently, it makes a low use of evaluation for decision-making and evidence-based citizen action. There is no broad awareness about evaluations as a public good, while the actual access to this information is limited and eventual. The complexity of reports that make them incomprehensible to citizens in general, low demand, and insufficient competencies of civil society in general to make use of available information must be added. In that context it is important to mention the role that REDMEBOL has been playing since 2017, focusing much of its efforts to promote the culture and institutionalisation of evaluation in the country, from an inclusive and transformative approach. Also noteworthy is the creation of the EJE Bolivia.

As for the academy, both documentary review and the interviews carried out revealed that no formal academic offers are identified in the country in order to professionalise M&E, that is, there is no academic training granting a higher diploma in the field of evaluation. The interviewees mention that the majority of projects in which consultants work assume that the latter should already have been trained in evaluation and very few projects consider including the development of skills or competencies in this area.

## LIST OF ABBREVIATIONS

CEBEM	Centro Boliviano de Estudios Multidisciplinarios (Bolivian Centre for Multidisciplinary Studies)
-------	---

CEUB	Comité Ejecutivo de la Universidad Boliviana (Bolivian University Executive Committee)
EJE	Equipo de Jóvenes Evaluadores/as de Bolivia (Team of Young Evaluators of Bolivia)
Evaluar	Red Argentina de Evaluación (Argentina Evaluation Network)
IDB	Banco Interamericano de Desarrollo (Inter-American Development Bank)
INE	Instituto Nacional de Estadística (National Statistics Institute)
MDP	Ministerio de la Presidencia (Ministry of the Presidency)
MDPyEP	Ministerio de Desarrollo Productivo y Economía Plural (Ministry of Productive Development and Plural Economy)
MDRyT	Ministerio de Desarrollo Rural y Tierras (Ministry of Rural Development and Land)
MPD	Ministerio de Planificación del Desarrollo (Ministry of Development Planning)
M&E	Monitoreo y Evaluación (Monitoring and Evaluation)
PDES	Plan de Desarrollo Económico y Social (Economic and Social Development Plan)
POA	Plan Operativo Anual (Annual Operating Plan)
PSDI	Plan Sectorial de Desarrollo Integral para <i>Vivir Bien</i> (Integral Development Sector Plan to Live Well)
REDMEBOL	Red de Monitoreo y Evaluación de Bolivia (Network for Monitoring and Evaluation of Bolivia)
ReLAC	Red de Seguimiento, Evaluación y Sistematización de Latinoamérica y el Caribe (Network for Monitoring, Evaluation and Systematisation of Latin America and the Caribbean)
RIEPP	Red Internacional de Evaluación de Políticas Públicas (International Network for the Evaluation of Public Policies)
SEGIP	Sistema General de Identificación Personal (General Personal Identification System)
SPIE	Sistema de Planificación Integral del Estado Plurinacional (Plurinational State Integral Planning System)

UDAPE	Unidad de Análisis de Políticas Económicas (Economic Policy Analysis Unit)
UNESCO	United Nations Educational, Scientific and Cultural Organisation (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura)

## LIST OF INTERVIEWS

<i>No</i>	<i>NAME</i>	<i>SECTOR</i>	<i>POSITION</i>	<i>DATE</i>
1	Cesar Ayala Gonzales	Public	Head of Strategic Planning—General Directorate of Planning	27-06-2018
2	Milton Málaga Arteaga	Public	Head of Strategic Planning—General Directorate of Planning (interviewed in a particular and non-institutional way)	05-07-2018
3	Diana Urioste	Cooperation	Director	09-07-2018
4	Rodolfo Soriano	Independent	Consultant, evaluator	10-07-2018
5	Walter Ferreira	Public	Consultant	15-07-2018
6	Roland Pardo	Public	Director of Social Policies	27-07-2018
7	Mike Gemio Pérez	Public	Executive Director General, FONABOSQUE. (interviewed in a particular and non-institutional way)	31-07-2018
8	Ivonne Farah	Academia	Economist, professor and researcher CIDES UMSA	26-06-2018

## REFERENCES

- Asamblea Legislativa Plurinacional. (2010). *Ley de la educación ‘Avelino Siñani - Elizardo Pérez’ No. 070*. [http://www.minedu.gob.bo/files/documentos-normativos/leyes/LEY\\_070\\_AVELINO\\_SINANI\\_ELIZARDO\\_PEREZ.pdf](http://www.minedu.gob.bo/files/documentos-normativos/leyes/LEY_070_AVELINO_SINANI_ELIZARDO_PEREZ.pdf). Accessed on 13 October 2020.

- CEUB. (2003). *Procesos de evaluación externa y acreditación en universidades bolivianas*, Bolivia. [http://www.ceub.edu.bo/academica/documentos/normas/21\\_Reglamento\\_Eval\\_y\\_Acred.pdf](http://www.ceub.edu.bo/academica/documentos/normas/21_Reglamento_Eval_y_Acred.pdf). Accessed on 13 October 2020.
- Diario Página Siete. (2015) *¿Por qué en Bolivia no se evalúa la educación? Interview conducted with Luis Oporto*. <https://www.paginasiete.bo/ideas/2015/8/30/por-bolivia-evalua-educacion-68084.html>. Accessed on 13 October 2020.
- El honorable Congreso Nacional. (1990). *Ley N° 1178. De administración y control gubernamentales*. <https://www.lexivox.org/norms/BO-L-1178.html>. Accessed on 13 October 2020.
- Estado Plurinacional de Bolivia. (2009a). *Constitución política del estado*. Plurinacional de Bolivia. [https://www.oas.org/dil/esp/Constitucion\\_Bolivia.pdf](https://www.oas.org/dil/esp/Constitucion_Bolivia.pdf). Accessed on 13 October 2020.
- Estado Plurinacional de Bolivia. (2009b). *Decreto supremo No. 29894. Para establecer la estructura organizativa del órgano ejecutivo del Estado Plurinacional*. Estado Plurinacional de Bolivia. [www.oas.org/juridico/PDFs/mesicic3\\_blv\\_ds29894.pdf](http://www.oas.org/juridico/PDFs/mesicic3_blv_ds29894.pdf). Accessed on 14 October 2020.
- Estado Plurinacional de Bolivia. (2012a). *Reglamento general de la cámara de diputados*. Asamblea legislativa plurinacional. [www.diputados.bo/sites/default/files/reglamentogral.pdf](http://www.diputados.bo/sites/default/files/reglamentogral.pdf). Accessed on 13 October 2020.
- Estado Plurinacional de Bolivia. (2012b). *Reglamento general de la cámara de senadores*. Asamblea legislativa plurinacional. [senado.gob.bo/file/67/download?token=OwEjkjCz](http://senado.gob.bo/file/67/download?token=OwEjkjCz). Accessed on 13 October 2020.
- Estado Plurinacional de Bolivia. (2016). *Ley No. 777. Para la creación de sistema de planificación integral del estado - SPIE*. Estado Plurinacional de Bolivia. <https://www.lexivox.org/norms/BO-L-N777.xhtml>. Accessed on 15 October 2020.
- Ministerio de Planificación del Desarrollo. (2016). *Lineamientos metodológicos para la formulación de planes sectoriales de desarrollo integral para 'Vivir Bien' PSDI*. [http://www.planificacion.gob.bo/uploads/PSDI\\_FINAL.pdf](http://www.planificacion.gob.bo/uploads/PSDI_FINAL.pdf). Accessed on 13 October 2020.
- Ministerio de Salud. (2012). *Resolución Ministerial No. 0092. Para la aprobación del Manual de Organización y Funciones del Ministerio de Salud del Estado Plurinacional de Bolivia*.
- REDMEBOL. (2017). *Boletín N° 3, Gestión 2017*. La Paz, Bolivia. <https://www.evalpartners.org/sites/default/files/documents/p2p/2017/June%202017%20-%20REDMEBOL%20Newsletter.pdf>. Accessed on 25 October 2020.
- REDMEBOL. (2018). *Institucionalización de la Evaluación en Bolivia*. Printed Format. REDMEBOL.
- UDAPE. (2014). *Manual de organización y funciones (MOF)*. <http://www.udape.gob.bo/docs/MOF.pdf>. Accessed on 13 October 2020.



## Evaluation in Brazil

*Verena Jacques Dolabella*

### INTRODUCTION

The universe of evaluation in Brazil is very plural and multifaceted. Many sectors within the society use the concept with different meanings, with different practices and with quite different levels of development. Addressing this issue in its entirety is certainly a challenge. There is, therefore, a need to attempt to bring up this multiplicity: government, civil society, consultants, international agencies, private foundations, academy, researchers, for one same job. Convergence can be a way of promoting dialogue between these actors and possibly build a more integrated future for evaluation in the country. Our starting point will be the normative aspects and the advances that the country has made in this direction, then we will go through more practical aspects within the governmental and civil society spheres regarding the use of the evaluations performed. Finally, we will deal with professional initiatives in the country.

Considering the institutionalisation of evaluation under the normative aspect may be the beginning of the analysis, but as we will see, it is incipient for understanding the practice and aspects transversal to the

---

V. J. Dolabella (✉)  
Collaborative Impact, Rio de Janeiro, Brazil  
e-mail: [verena@collabimpact.org](mailto:verena@collabimpact.org)

evaluation in the country. In an attempt to broaden the analysis spectrum, we had conversations with the main evaluators in the country in line with a historical profile of their performance, especially during the years 2003 to 2011. We will see how several political aspects are deeply connected to the real practice of evaluation. Therefore, an analysis of the subject is not possible only from its products (the evaluations themselves) but also from the entire structure that enables the collection and dissemination of data in the country, often threatened according to government ideology (data collection and dissemination systems have been questioned by the Bolsonaro government, for example). We will, therefore, speak about the maturity of this Brazilian national democracy, which is still recent, with just over 30 years. This article considered the legislation, government regulations and programmes in force as of May 2020. Any changes after this period had not been considered or analysed in the context of this publication.

In addition, reflection on the role of evaluation and the its development as discipline and as a profession in the country is becoming increasingly fundamental. There are fertile paths to be followed, thinking about consolidating learning and improving public policies in the country.

## GENERAL COUNTRY OVERVIEW

The country's social and economic-political structure is the starting point for a broader analysis of the Brazilian context. The country is constituted as a federative Republic, with 27 states, including the administrative headquarters, the Federal District, and 5570 municipalities throughout the national territory. There is state (and district) autonomy for the creation of laws and public policies, subject to the Federal Constitution. Such autonomy is reflected in the actions of institutionalisation of the evaluation, where significant advances can be noted in some states, while in others the administrative structure still needs to be developed in order to approach evaluation possibilities. This lack of unity in relation to the administrative and legal system makes the analysis of the country, which has continental dimensions, more complex and therefore requiring a very specific case-by-case analysis.

The term 'coalition presidentialism', coined by political scientist Sérgio Abranches (1988), is still used today to designate the Brazilian political system, which combines presidentialism, proportional representation and multiparty politics, governed by large political party coalitions. Brazil has



more than 33 political parties<sup>1</sup> that organise themselves into coalitions to ensure governability for the elected president.

The government controls legislative production, and this control is the result of the interaction between agenda power and support from the majority. The majority brought together solely by party coalition. Nothing quite different from what happens in parliamentary governments. (Limongi, 2006, p. 25)

Despite the multiplicity of political parties, they are not central to the country's political construction, playing a passive and secondary role, far from representing different political proposals and ideas. This fact, concretely, makes these parties articulate votes in exchange for positions and elections of deputies and senators or even indirectly of mayors and governors (Matusceli, 2010). In the legislative sphere, the thesis of the balance of powers between the legislative and the executive has been questioned a lot, since there is no dispute of actual proposals with the executive. The exchange of positions and elections is the predominant logic in the relationship between the powers, emptying the political debate and the real meaning of the democratic structure. This structure and political logic also affect the institutionalisation of evaluation environment. It shows that it takes more than an organised group, with technically and politically well-structured proposals. The Brazilian context demands more than good proposals for laws and policies to be approved. The political game has a logic linked to the old power structures, the construction of arrangements interested not in the common good, but in financial and political gain. Another point about this logic is that it does not benefit the creation of a national policy, but it potentialises initiatives in a more limited context, such as those of the state, where the actors relate more closely around the agendas.

The figure of the president of the republic is central to the population (it is not the parties or the legislature), and his decision-making capacity is more linked to state bureaucracy and less aligned with the representativeness of his actions. This arrangement is defined as hyperpresidentialism by some authors (such as Matuscelli, 2010 and Torres, 1996), and points to the fragility of democracy in Brazil, where in practice there is no inclusion

<sup>1</sup> According to the website of the Superior Electoral Court at <http://www.tse.jus.br/partidos/partidos-politicos/registrados-no-tse>. Accessed in May 2020.

and participation of all social strata in the political system. Political personalism in the figure of the president is strong and ends up hiding other forces in the political sphere, still strongly dominated by the country's economic elites.

From an economic point of view, a country of continental dimensions has a strong economy in absolute numbers—the country is considered the ninth world economy by the 2018 IMF report.<sup>2</sup> But it is necessary to expand the analysis to the point of view of the internal distribution of wealth: However, the territorial inequality in the country's wealth generation is still significant. According to the study by the Institute of Applied Economic Research (IPEA),<sup>3</sup> only 1% of the wealthiest municipalities accounted for 47% of the Brazilian Gross Domestic Product (GDP). On the other hand, 70% of the poorest cities represent 14.7% of GDP. States and municipalities have different collections, and inequality is also regional: south and southeast with economic centrality, while other states, the majority of the general population, are still facing problems such as hunger, lack of basic sanitation and essential services. From an evaluative point of view, the inequality of infrastructure and social services also makes the structuring of a national evaluation policy or law complex: each region with a different diagnosis and demand for actions makes it difficult for national policies to point out generalisations.

Organised civil society and social movements became stronger since the process of re-democratisation. The 1988 Constitution (BRASIL. [Constituição (1988)]) provides an opening for social participation in policy formulation and monitoring processes with the Law Councils. Some political agendas emerged from the social struggle, as for the health issue, with the defense of a single and unified health system (SUS), or in the anti-asylum struggle movement, with policies to combat hunger and affirmative rights policies. Despite great achievements, the performance of these movements is still marginal,<sup>4</sup> so to speak, constantly threatened and without guarantee and security for action.

<sup>2</sup> For more information: <https://www.imf.org/en/Publications/WEO/Issues/2018/09/24/world-economic-outlook-october-2018>. Accessed in May 2020.

<sup>3</sup> For more information, see: Income Inequality in the Brazilian Territory at <http://repositorio.ipea.gov.br/handle/11058/5291>. Accessed in May 2020.

<sup>4</sup> Social movements for land and housing, for example, are often criminalised. See: Grzybowski 1991.

The media in the country has always been a portrait of the political organisation, where economic elites have control and monopoly, and prevented the broad participation of society.<sup>5,6</sup> The main communication groups in the country belong to elite families and still have a great influence on the knowledge of the population in general. It is undeniable the growth of independent journalism in the country provided by the internet and alternative communication channels, still with funding difficulties. This communicational openness is promising greater participation of society in politics, greater access to information and debate of ideas. The independent media has suffered constant attacks intending to make it unfeasible and discredited. The internet has also been occupied by criminal groups, producing fake news, uttering hatred and waging attacks against democratic institutions.

Federal universities became centres of knowledge; despite the growing emergence of the private sector, cutting edge knowledge is still concentrated in public universities with established researchers and research tradition. Universities, as well as social movements, are increasingly assuming a role of resistance: with cut investments, their production is threatened and their relevance questioned by those in power, in the current context.

Other important social actors in the country are international agencies, foundations and civil society organisations. The country already had sizeable international financing for projects in the social area and the effects of these investments extend to the present: a generation of researchers and managers was formed based on international cooperation and on the humanitarian ideas of these agencies. With the country's economic growth, especially in the first decades of the 2000s, this investment migrated to countries in the African continent, and a culture of private social investment, still on the rise, is beginning to emerge in the country. As we will see, private social investment has created a fertile environment for the debate on evaluation. Its political independence allows more possibilities for learning and methodological debate, while on the

<sup>5</sup> Five families control half of the 50 media outlets with the largest audience in Brazil. The conclusion is from the Media Ownership Monitor survey, funded by the German government and carried out jointly by the Brazilian NGO, *Intervozes*, and Reporters Without Borders, based in France.

<sup>6</sup> For more information see Sodré 1999.

other hand, the use of evaluations is still an evolving process, as we will see.

From a historical point of view, an analysis of evaluative thinking in the country should start in the post-1988 constitution period, with the process of re-democratisation of the country. The constitution inaugurates a period of conception of public policies based on universality, on the broad access to health, assistance and education services. The consolidation of policies with an increase in scope and scale requires more investment in the state apparatus, which had its technical staff significantly increasing its analytical capacity and improving management tools (Jannuzzi, 2020). There are also mechanisms for social participation such as councils and laws for participatory budgeting in some municipalities.

Along with the development of social policies, the Brazilian state was also influenced by the neoliberal wave experienced by the United States in the 1960s, New Public Management, influencing countries to apply management focused on results and evidence. International agencies such as the World Bank and the Inter-American Development Bank (IDB) are central to this process and in the creation of an evaluation culture focused on evidence and methodologies in the economic area.

The reforms of the 1990s, with the decentralisation of administration, the increased demand for the quality of public services and the globalised financial market started to demand from governments processes and institutions to coordinate the decision-making process to guarantee political coherence and alignment between strategy, budget and management. (Oliveira, 2019, p. 14).

Another element that historically favoured the evaluative context was the recent fiscal adjustment, a guideline also updated by international organisations.<sup>7</sup> Such adjustment was necessary in a context of economic depression—Brazil has in 2014 a drop in revenue and low GDP growth, decreasing the payment capacity of officials and ministries. The crisis in the states is even deeper and many reach a situation of economic collapse. The adjustment demands ad hoc arrangements to evaluate and plan future

<sup>7</sup> See World Bank's Country Partnership Strategy for Brazil at <https://www.worldbank.org/en/news/press-release/2017/07/13/brazil-world-bank-group-presents-country-partnership-strategy>. Accessed in May 2020.

actions, thinking not only in the control of spending but also in the evaluation of the effectiveness of public spending with policies that have an impact on people's lives. We will see in the following chapter on the creation of the first Public Policy Monitoring Evaluation Council, the CMAP.

### *Institutional Structure and Processes Normative Aspects*

From a normative point of view, Brazil has some elements that point to an institutionalisation of the evaluation in the country, although there are no structured national laws about this matter.<sup>8</sup> In order to better understand the normative bases, we will carry out an analysis according to table 4.1 below, separating laws and regulations from the (i) national and (ii) sectoral perspective, (iii) policies and strategies and (iv) administrative regulation. In addition, we will analyse the institutionalisation of the evaluation in the parliament and we will approach a state case study

<sup>8</sup> In March 2021 it was approved an important change in the normative status in Brazil, giving the evaluation of public policies constitutional status. The Constitutional Amendment N° 109 amended Articles 37 and 165 of the Constitution. A new paragraph (16) of Article 37 was included providing that: "The organs and entities of the public administration, individually or jointly, shall conduct evaluation of public policies, including disclosure of the object to be evaluated and the results achieved, according to the law." The main purpose of the Constitutional Amendment was to allow the payment of emergency aid in Covid-19 context throughout 2021 and create instruments for future tax adjustments, institutionalising a process which was already in course since 2016 (Amendment N° 95), and compensating part of the extra expenses with the aid. The institutional development of monitoring and evaluation is, therefore, in dialogue with this matter. After the approval of the Constitutional Amendment, it is possible to observe updates of evaluation reports on CMAP's website as well as the publication of an Annual Report of Evaluation of Public Policies (<https://www.gov.br/fazenda/pt-br/orgaos/secretaria-de-avaliacao-planejamento-energia-e-loteria/documentos/cmap/relatorio-anual-de-avaliacao-de-politicas-publicas/view>, accessed on 12 October 2021) by the new Secretariat of Evaluation, Planning, Energy and Lottery, covering CMAP's work since its creation with the results of 16 evaluations of public policies in different sectors. The report brings information about the ongoing process of institutionalisation of M&E in the country, with CMAP in the centre and IPEA, ENAP and IBGE as main support partners. CGU is the governmental organ responsible for monitoring the recommendations of the evaluations. Also, the report indicates the challenge of systematising the evaluations results with the country's financial and budgetary process (article 165, paragraph 16 of the Amendment N° 109). Future actions are planned in the direction of further regulating the Constitutional Amendment N° 109. A deeper analysis of the impact of this Amendment in the M&E context is necessary, which has not been conducted in the context of this chapter.

**Table 4.1** Normative basis on Monitoring and Evaluation in Brazil (Author’s development)

<i>National Law or Regulation</i>	<i>Sectorial Law or Regulation</i>	<i>Policies or Strategies</i>	<i>State Law or Regulation</i>	<i>Administrative Regulations about Evaluation</i>
<b>Regulation:</b> Committees for Monitoring and Evaluation of Federal Public Policies (CMAP) Decree N° 9,834, of 12 June 2019	<b>Law:</b> Education (National Educational Plan, PNE) Art. 3 a 7 Law N° 13,005/2014	<b>Strategy:</b> Health (National Programme for improvement of quality and access to basic attention, PMAQ), Ordinance N° 1,645, of 2 October 2015	<b>State Law:</b> Espírito Santo (Public Policy Monitoring and Evaluation System, SIMAPP) State Law N° 10,744/2017	<b>Regulation:</b> Regulatory Impact Analysis for proposals of normative acts, Decree N° 10,411 of 30 June 2020
<b>Law:</b> Pluriannual Plan Approval Law (PPA): (valid only for the mentioned period)	<b>Law:</b> Education (National System of Higher Education Evaluation, SINAES) Law N° 10,861, DE 14 DE ABRIL DE 2004	<b>Strategy:</b> Economy Reasearch (Instituto de pesquisa em Economia Aplicada IPEA) Law N° 8,029, of 1990	<b>State Decree:</b> Ceará: (State Fund for Combating Poverty, FECOP and its Monitoring and Evaluation System) Art. 30 Decree N° 29,910, of 29/09/2009	<b>Resolution:</b> Legislative Consultancy making recommendations for performance of activities and the improvement of the legislative technique, Art. 2° Resolução da Câmara dos Deputados (Resolution of Deputy Chamber) N° 48, 1993

(continued)

based on the first case of a specific law that regulates the evaluation of policies in a state. Table 4.1 summarises the following analysis:

*National Laws and Regulations*

There is still no national law establishing parameters and guidelines for the evaluation of policies in an institutionalised manner. There are national

**Table 4.1** (continued)

<i>National Law or Regulation</i>	<i>Sectorial Law or Regulation</i>	<i>Policies or Strategies</i>	<i>State Law or Regulation</i>	<i>Administrative Regulations about Evaluation</i>
<b>Law:</b> Fiscal Court Unit, TCU (Budget Guidelines Law, LDO) Role of TCU auditing PPA and Reporting on Government Policy and Programme Monitoring Art. 24 Law N° 13,707/2018	<b>Regulation:</b> Social Development (Secretary for Evaluation and Information Management, SAGI) Ordinance N° 329, of 11 October 2006. Revoked by Ordinance Ministry of Social Development (MDS) N° 2,227/2018	<b>Strategy:</b> M&E Training (Public Management National School, ENAP) Ordinance Ministry of education (MEC) N° 660/201	-	-

regulations that point out recommendations and establish processes within institutions and agencies, for example:

### **Multiannual Plan**

The Multiannual Plan (PPA), in Brazil, provided in article 165 of the Federal Constitution is a medium-term plan, which establishes the guidelines, objectives and goals to be followed by the Federal, State and Municipal Government over four years. There are no specific laws that provide for the mandatory monitoring and evaluation of the PPA, it is noted, however, that the law that approves the plan in the last four federal administrations had a section that cites an evaluation report on the goals of the plan, which is subject to the government guidelines.

### **Federal Courts of Accounts—TCU**

From a budgetary point of view, the TCU is the body that acts in the external control of the PPA and of the two planning documents provided for in Brazil: Budget Guidelines Law and Annual Budget Law. There is also a trend within the TCU to start looking not only at budget execution but also at the effectiveness of implemented policies, for example:

the State Court of Accounts (TCE) of the state of *Ceará*, which through the *Plácido Castelo* Institute qualify public servants with focus on evaluation as a way of building public value. In addition, over the past three years the Court has begun to prepare the Report on Government Policy and Programme Monitoring in compliance with Article 124 of the Budget Guidelines Law (LDO) Law 13,707/2018, which analyses a sample of federal public policies and makes important recommendations on monitoring and evaluation, as we follow in more detail in the next chapter.

### **Policy Monitoring and Evaluation Council—CMAP**

One of the initiatives to contain expenses during the Dilma (2014–16) administration was the establishment of an Interministerial Working Group for Monitoring Public Spending (GTAG) of the main federal Programmes with the participation of 4 ministries (Civil House, Planning, Finance and the Comptroller-General of the Federal Union (CGU)). This experience was undoubtedly a political opportunity to expand evaluation mechanisms, culminating in the establishment of the CMAP, which was first named as a Committee, now named as a Council. Despite the focus on expenditure restraint, it was a unique moment of inter-ministerial articulation in the country, where executive secretaries presented their policies, identified flaws in the design, management and implementation and brought the most articulate information for the ministers to decide. This structure could be articulated due to several factors: presence of key factors in the design of monitoring and evaluation in the country, opportunity for agenda setting and external demand for fiscal balance (strong argument for these factors to coincide). This period is described in detail in the work of Patricia Oliveira, an important reading to understand this moment in the country's history. Some interesting points for analysis of this period:

- Motivated by the focus on cutting costs, establishment of an Interministerial Working Group for Monitoring Public Spending (GTAG) at 2015. Fiscal tightening, decision-making for resource reallocation;
- The exchange of information between executive secretaries and CGU and Treasury control bodies, for example: 'More information and more decision-making capacity were generated,' 'Greater Foundation'



- Some subgroups had the opportunity to address the management of the policies that were being evaluated;
- Subjective and conflicting objectives, despite the emergence of the evaluation agenda, the main focus was on reducing spending;
- Identification of management problems in selected policies (fraud in programmes, inefficiencies in deliveries, poorly designed policies, lack of indicators and monitoring data). Anyway, poor quality of public spending;
- Moment before the creation of the CMAP (which repeated the GTAG model), the established working group came close to structuring an evaluation of public policies system with the proposal of the System for Monitoring and Evaluation of Public Policies and Public Expenditure Review (SIAPRE)—a system that according to its creators “Structured and much more effective for you to really fight bad public policy. Effectively able to change the modus operandi of public policies.” Until today awaiting a president willing to sign it;
- CMAP structure was also a controversial issue (IPEA, ENAP, *Instituto Brasileiro de Geografia e Estatística* (IBGE) and Government Budgeting Management School (ESAF) —as consulting bodies) Civil House, Ministry of Finance, Ministry of Planning, CGU;
- Ambition to influence budgeting;
- Meaning of policy analysis and inter-ministerial integration; and
- Only possible due to the political capacity of the actors of the moment to set an agenda.

As the contextual factors matter in the unfolding of the institutionalisation of the evaluation, this period was interrupted by a fracture in the political system: the ousting of the president via impeachment. The pressure of the demand for fiscal adjustments, combined with street manifestations and a process in the sphere of fiscal responsibility and political articulation of his opponents, culminated in the sudden change of government where Dilma Rousseff’s vice-president (Michel Temer) took over. As a result, there was an exchange of ministers and the work of CMAP, which had already begun at the time, had started to be revised by Michel Temer’s Team.<sup>9</sup>

<sup>9</sup> For more information, see the report: Federal Public Policy Monitoring and Evaluation Committee Studies and Proposals [http://www.ipea.gov.br/portal/images/stories/PDFs/livros/livros/181127\\_comite\\_de\\_monitoramento\\_inicio.pdf](http://www.ipea.gov.br/portal/images/stories/PDFs/livros/livros/181127_comite_de_monitoramento_inicio.pdf). Accessed in May 2020.

From the normative point of view, the technical team that was part of CMAP arrangement before the impeachment managed to move forward and CMAP was instituted by decree (Decree No. 9834, 12 June 2019) gaining the status of council and assuming itself as an advisory body and not deliberative as the Committee's first plan was. In addition, the decree establishes two working 'Committees' for the council, one related to the subsidies policy (CMAS—Monitoring and Evaluation Committee of Union Subsidies) and the other to the direct expenditure (CMAG—Monitoring and Evaluation Committee of Union Direct Expenditure). The decree defines CMAP as:

Federal Public Policy Monitoring and Evaluation Council - CMAP, with the aim of: I - improving public policies, Programmes, and actions of the federal Executive Branch so that they achieve better results; and II - improving the allocation of resources and improve the quality of public spending. (Art.1, Decree No. 9834, 12 June 2019)

Despite the normative achievement and some important technical initiatives,<sup>10</sup> the discontinuity and dispersion of the technical staff with the transition of the governments drained the Council's political strength—if in the previous two moments it started from a government demand, it currently has not the same function within the government. An important point was the lack of inclusion of civil society in this process, which was configured as an initiative by the Government Centre.<sup>11</sup> There is the hypothesis that the guarantee of instruments that involve civil society, could also have guaranteed the strength of this group. But it remains only a hypothesis, since the actors who had common interests moved away and this political agenda is no longer a priority. Other spheres recognise the importance of this institutional achievement, such as the TCU for example:

<sup>10</sup> Some important publications for these periods:

- Guia de Análise Ex Post (BRAZIL, Casa Civil et al. 2018a), prepared by the Civil House of the Presidency of the Republic, by the Ministry of Finance (at the time Ministry of Finance and Ministry of Planning, Development and Management), by the CGU (at the time Ministry of Transparency and General Controllorship by the Union) and by the IPEA, with the participation of other agencies;
- BRAZIL, Casa Civil et al. 2018a & 2018b.

<sup>11</sup> Term well detailed by Oliveira 2019, p. 23.

Finally, it is important to point out that there are recent initiatives of the federal government, such as the institutionalisation of the CMAP, through Decree 9,834/2019. The monitoring and stimulation of the effective performance of this council and its committees can contribute to the improvement of the government's capacity to monitor and evaluate public policies. (BRAZIL, The Federal Court of Accounts, Report on Government Policy and Programme Monitoring (RePP) 2019, p. 31, item 129)

The most recent action of the Council was designated by the CMAS creating a Technical Group to draw up a proposal for a Governance Model for Union Subsidies.<sup>12</sup>

The Council had their second ordinary meeting by the 13 March 2020 deliberating about the design of a fishing policy and a proposed methodology on public policy selection criteria to be assessed from the 2020–2023 PPA Programmes.

#### *Sectoral Laws and Regulations*

#### **Secretary for Evaluation and Information Management—SAGI (Social Development)**

The Secretary for Evaluation and Information Management linked to the Ministry of Social Development through administrative regulation (Ordinance No. 329, of 11 October 2006 and revoked by MDS Ordinance No. 2,227/2018) is the main sphere of monitoring and evaluation of social policies located within a ministry. Historically, it had great importance during Lula and Dilma governments, where social policies (such as *Bolsa Família* and *Programa Brasil sem Miséria*) were the focus of the government programme. The secretariat was the first monitoring and evaluation policy initiative in the country, and as we will see, it played a key role in fostering and institutionalising evaluation in Brazil. Its challenges were great: to integrate data on social policies implemented by various secretariats (social assistance, education, health) and to establish an evaluation culture on the part of the public management structure and the civil servants.

<sup>12</sup> CMAS Resolution, N° 01, 30 April 2019 at <https://www.gov.br/economia/pt-br/aceso-a-informacao/participacao-social/conselhos-e-orgaos-colegiados/cmap/publicacoes/atas-e-resolucoes-1/resolucoes/99-resolucao-cmas-no-01>. Accessed in June 2020.

SAGI then needed a policy of convincing managers—its internal clients—that the information produced by monitoring and evaluation would make it possible not only to improve the performance of the programmes, but also to verify whether the expected results were being achieved. Over time, SAGI's role became clearer, recognised and legitimised by other secretariats. (Vaitsman et al., 2006, p. 17)

The training of the team was also an important aspect in the constitution of the Secretariat, formed by a very young team, some of them fresh from graduation or master's degree, or still attending the master's degree. Lack of a tradition in the area of evaluation and monitoring of social programmes in the public sector also emphasised the need for training of the team. SAGI has, over little more than two years of existence of the MDS, invested on the participation of the team in different courses of evaluation of social policies and programmes available in the country and even abroad. In addition, SAGI itself financed the course on Evaluation of Social Programmes, in partnership with other institutions and universities. It also encouraged the participation of technicians in national and international seminars and congresses (Vaitsman et al., 2006).

Its resources have come from various sources: National Treasury, World Bank and IDB loans. Part of the Treasury's resources, that is, the national budget allocated to the MDS annually and approved by the National Congress, was executed through cooperation projects with multilateral organisations such as the United Nations Educational, Scientific and Cultural Organisation (Vaitsman et al., 2006).

Two systems were set up within SAGI: one for monitoring and one for evaluation. The monitoring system went through the challenge of structuring the database from a collection of data from different municipalities and secretariats, passing through instances responsible for resource distribution, such as *Caixa Econômica* Federal Bank at the Bolsa Família programme. It is worth noting that many municipalities did not have access to the Internet and the collection of these data was still done manually. SAGI developed a tool for the database and was responsible for developing the analysis indicators. Other processes were foreseen during the implementation of the system: the creation of a digital glossary, a work group involving servers from various areas, among other initiatives that legitimised the processes within the different teams, consolidating the importance of monitoring for social policies (Vaitsman et al., 2006).

The evaluation system prioritised mainly external evaluations, through a Term of Reference setting out the priority policies. The criteria for this choice depended on the equation: resources available for the evaluation, structured data, objective, and time for execution of the evaluation.

The dissemination of data and studies was structured through several publications, such as *Cadernos de Estudos de Desenvolvimento Social* and *SAGI em Foco*.<sup>13</sup>

Currently, SAGI's function has been altered by MDS Ordinance no. 2,227/2018,<sup>14</sup> and its structure has been undone. It is clear that SAGI's strategic position has changed politically, no longer having the relevance it had in previous governments. The number of studies published annually is a good indicator to understand this loss of relevance in the execution of policy evaluation: from 2010 to 2014, 81 studies of public policy analysis were made while from 2015 to 2018 only 13.<sup>15</sup>

### National Education Plan—PNE (Education)

Article 124 of the Federal Constitution and embodied in Law 13,005/2014 deals with the National Education Plan—PNE, which can be seen as good sector practice to be replicated. The law objectively establishes guidelines and goals to be achieved, thus guiding on issues related to the monitoring and evaluation of that policy:

Art. 5 The execution of the PNE and the fulfillment of its goals will be the object of continuous monitoring and periodic evaluations, carried out by the following instances:

I - MEC;

II - Education Commission of the House of Representatives and Education, Culture and Sport Commission of the Federal Senate;

III - National Council of Education;

IV - National Education Forum.

§ Paragraph 1 It is also incumbent upon the bodies mentioned in the caput:

<sup>13</sup> Available on SAGI's website <https://aplicacoes.mds.gov.br/sagi/portal/index.php?grupo=164>. Accessed in June 2020.

<sup>14</sup> The ordinance, among other actions, limits SAGI's capacity for analysis as the first paragraph of article 9º stands “§1 SAGI will not comment on the criteria of opportunity or convenience of the MDS Secretariats” (Art 9º Ordinance N° 2,227, 6 June 2018).

<sup>15</sup> Analysis of data available at: <https://aplicacoes.mds.gov.br/sagi/portal/index.php?grupo=182>. Accessed in May 2020.

I - to disseminate the results of the monitoring and evaluations on the respective institutional Internet sites;

II - to analyse and propose public policies to ensure the implementation of strategies and the achievement of goals;

III - to analyse and propose the revision of the percentage of public investment in education.

§ Paragraph 2 Every 2 (two) years, throughout the period of validity of this PNE, the National Institute of Educational Studies and Research *Anísio Teixeira* (INEP) will publish studies to assess the evolution in the fulfillment of the goals established in the Annex of this Law, with information organised by federal entity and consolidated at the national level, having as reference the studies and researches referred to in Article 4, without prejudice to other sources and relevant information.

§ Paragraph 3 The progressive goal of public investment in education shall be evaluated in the fourth year of effectiveness of the PNE and may be expanded by law to meet the financial needs of meeting the other goals. (Art. 5 of Law 13,005/2014)

As one can see, the law that instituted the policy has explained its guidelines, goals, governance structure, as well as its model for monitoring, evaluation and accountability. These actions made it possible to know the scope of the policy and monitor its performance through the results of biannual evaluations undertaken by INEP.<sup>16</sup> In addition, as indicated in Article 5º, mechanisms for social participation in the evaluation of the policy, such as the National Education Forum, are planned.

Another law in the education Sector establishes the National System of Higher Education Evaluation (Law N° 10,861, 14 April 2004), with the aim of ensuring a national process of evaluation of higher education institutions, undergraduate courses and academic performance of their students. This system is coordinated within the state's educational sector and has established instruments and tests for students and universities.

### *Policies and Strategies*

#### **The National Programme for Improvement of Basic Care Access and Quality (PMAQ)**

This is an example of a monitoring and evaluation system: created by Ordinance N° 1,645, October 2015, as a part of the Brazilian Unified

<sup>16</sup> Available at: <http://portal.inep.gov.br/web/guest/dados/monitoramento-do-pne/relatorios-de-monitoramento>. Accessed in June 2020.

Health System—SUS for massive data collection to widen the access to and the quality of health services in Brazil. The PMAQ was established in 2012, and one of its strategies is linking the quality of the services to the transfer of funds. The evaluation model chosen had a national scope and was structured based on a great number of indicators, training the team and the public manager for the service monitoring and evaluation, for which they are also liable. According to several experts in health evaluation (see Silva et al., 2015), the programme assumes that the evaluation initiative is a way of carrying out ongoing improvement of the service and the consolidation of a favourable environment to train on basic care.

### **Collecting and Databases**

Databases and the entire input system (statistics and information) for monitoring of public policies are a key part of this puzzle, such as the IBGE Census (1991, 2000, 2010) and more detailed surveys such as the National Household Sample Survey. The IBGE Census and data from the INEP or Data SUS (when thinking about education and health respectively) were the main diagnostic mechanism for expanding the scope and scale of social policies. The development of databases generated a growing demand for data, improving academic research capacity and the environment for evaluation. This structure, so important for the consolidation of scientific research in the country, has been emptied by the current government (2020), which makes less and less data available to the population. During some interviews conducted, many researchers claimed insufficient data to continue monitoring policies, especially data related to increasing poverty and inequality in the country.

### **The Applied Economics Research Institute (IPEA)**

The Institute conducts studies on government policies and programmes with technical staff and diverse methodologies. It is a federal public foundation linked to the Ministry of Economy by the IPEA Law N° 8,029, de 1990 and its research activities provide technical and institutional support to government actions for the formulation and reformulation of Brazilian public policies and development programmes. IPEA's work is made available to society through numerous and regular electronic publications, printed publications and events. Its technical capacity serves as a subsidy to Ministries, Secretariats and is the main technical instance of the CMAP. In the following chapter, we will make a brief analysis of the latest studies

conducted by the Institute and which are methodologically configured as evaluation of public policies.

### **National School of Public Administration (ENAP)**

This is an important training centre for public managers, which has been dealing with evaluation in their training. It currently holds a Professional Master's Degree in Public Policy Evaluation and Monitoring and concentrates some of the greatest specialists in government evaluation, many of them creators of CMAP. ENAP trained more than 400,000 employees in 2019 in its various courses, an important work of training the employees around technical capacity and knowledge of evaluation. There are also state government schools such as Fundação João Pinheiro in Minas Gerais e Instituto Jones in Espírito Santo.

### *Administrative Regulations About Evaluation*

The recent Decree No. 10,411 of 30 June 2020 regulates the Regulatory Impact Analysis, which must be carried out when there is a new normative proposition of general interest of economic agents or of users of the public services provided within the scope of their competence. The decree has effects on regulatory agencies and the Ministry of Economy, among others, and provides items that must be included in the Impact Analysis Report, such as (a) identification of the regulatory problem; (b) identification of the economic agents; (c) identification of the legal grounds that support the agency's action; (d) definition of objectives to be achieved; (e) description of the possible alternatives to the regulatory problems; (f) exposure of the possible impacts; (g) the alternatives identified; (h) considerations regarding the information and manifestations received for the analysis of the impact on eventual processes of social participation or other processes (i) mapping of international experience; (j) identification and definition of effects and risks resulting from the edition of the regulatory act; (k) comparison of the alternatives considered and (l) description of the suggested strategy.

### *Parliament*

The Parliamentary sphere doesn't have any laws concerning the evaluation field. The Committees, Parliamentary Fronts, could be institutional avenues for discussing the effects of legislation using evidence. Unfortunately, it usually doesn't happen. We still have legislative houses very



focused on the political game, using the vote of laws for political bargaining.

The Legislative Consulting is an important body producing evidence in the chamber. It is a consulting and institutional advisory body to deputies, composed by a multidisciplinary team of consultants selected in public competitive exams, divided into 22 thematic areas able to technically subsidise all stages of the legislative process and parliamentary activity. They are responsible for the production of Technical Notes. It is the practice that comes closest to a legislative evaluation. About 30 studies are published annually on the chamber's website.<sup>17</sup> The body competences are established by the internal legislation at the Art. 2º of the Resolution of the Deputy Chamber N° 48 (Câmara dos Deputados, 1993).

A recent practice organised by civil society organisations is the professionalisation of parliamentary offices with the hiring of a team of advisors responsible for building evidence and supporting the actions of the deputy. 'Legisl' is one organisation that has been working on this, with the aim of accelerating parliamentary mandates on three fronts: team composition, planning and management. There are also other innovative initiatives taking place in the chamber as collective and shared mandates (where a technical team integrates the mandate of an elected deputy in order to pluralise the political debate). These initiatives could increase the use and production of evaluations in legislative bodies, but we still can't see any empirical results.

### *State Case*

The state of Espírito Santo is the first national state to institutionalise evaluation in the planning and implementation cycle of public policies through an approved law. The SIMAPP was instituted by State Law 10,744/2017. The model of the system was designed over three years coordinated by the Jones dos Santos Neves Institute<sup>18</sup> with technical

<sup>17</sup> <https://www2.camara.leg.br/atividade-legislativa/estudos-e-notas-tecnicas/publicacoes-da-consultoria-legislativa/Estudos-e-notas-tecnicas>. Accessed in May 2020.

<sup>18</sup> Institute linked to the Secretariat of Economy and State Planning with the purpose of producing knowledge and subsidising public policies through the elaboration and implementation of studies, researches, plans, projects and organisation of statistical and georeferenced databases, in the state, regional and municipal spheres, focused on the socioeconomic development of Espírito Santo.

support from the Centre for Learning in Evaluation and Results for Lusophone Africa and Brazil (CLEAR LAB), the Institute for Learning and Research and the Ayrton Senna Institute. It's a partnership between state, international organisations, academia and civil society. It is interesting to think that this first experience of institutionalisation by law is part of a multidimensional arrangement, with allies from all fronts analysed by this article.

Some strategic premises guided this process, so that it would in fact be assimilated by public management:

- search for low-cost mechanisms: in line with the context of state tax adjustment;
- use of existing institutional mechanisms and structures: institutional support for the implementation of the system is based on existing institutions, such as technical implementation of evaluations with the Jones Institute, and the use of committees in the secretariats.
- training of public servants: the project was close to the secretariats' staff, so that they would be practices adopted in government routines;
- approximation with the academy: the academy is not only an important partner in the execution of evaluations with the technical staff, but also a network of innovation and methodological validation;
- Search for partnerships: internal and external partners that foster the field of evaluation were essential for an environment of cooperation and construction of a multisectoral policy.

The state of Ceará has also progressed in this agenda, mainly with the State Fund for Combating Poverty—FECOP, which, by means of administrative act, establishes procedures for the analysis and evaluation of projects sent to the Executive Management of FECOP, subsidised by the guidelines of the Centre for Data Analysis and Evaluation of Public Policies, of the Ceará Institute for Research and Economic Strategy. The decree<sup>19</sup> determines that the projects must contain, among other things, a proposal for the design of a grounded impact assessment.

<sup>19</sup> Ceará Decree N° 29,910, 29/09/2009.

### Other Proposals of Evaluation Legislation, Currently in Process

There are two important bills in the national congress, which if approved, would represent advances in the institutionalisation of evaluation in the country. It is important that civil society monitors these legislatures.

- Constitutional Amendment Proposal (PEC) 26/2017—It determines that the Executive, Legislative and Judiciary Powers shall maintain, in an integrated manner, a system of evaluation of public policies, with the objective of promoting the improvement of public management, which shall evaluate the economy, effectiveness, efficacy and efficiency of governmental actions; shall provide technical subsidies for the formulation of new public policies; shall observe the principle of periodicity; and, shall be exercised with the assistance of the Court of Auditors of the Union and of the organs integrating the internal control system of each Power.

Article 1 Articles 71 and 74 of the Federal Constitution shall come into force with the following amendments:

Article 71 § 5—Without prejudice to the provisions of Clause IV of the caput, the Court shall carry out, through operational audits, the long-term monitoring of public policies specified in the multi-annual plan, with the objective of evaluating their economy, effectiveness, efficacy and efficiency, as well as providing technical subsidies to formulating and executing agencies for their improvement. (NR)

Article 74 V—periodically evaluate, in the form of the law, the economy, effectiveness and efficiency of public policies, with the aim of providing technical subsidies for the improvement of government management and the formulation of new policies. (Excerpts from PEC 26/2017)

Status: Approved at CCJ, ready for plenary deliberation.

- Law Proposal (PL) 488, 2017—Federal Senate Bill

Amends the Law on Legislative Technique to provide that legislative proposals establishing public policies contain legislative impact assessment.

Status: Approved in the Senate, in process in the House.

## EVALUATION PRACTICE

The Report prepared in 2002 by the IDB in technical cooperation with IPEA presented some conclusions about Brazilian evaluative practice (BRAZIL, IPEA, 2002). Much progress has been made since then, but some points remain very pertinent: the report pointed out that the country has experimented with a wide variety of concepts and approaches concerning the evaluation function, especially in the context of the experience of planning, government management and developing social programmes at the federal level.

The Brazilian experience in this sector is broad and diversified but still considered insufficient and unsatisfactory. The evaluations of government programmes are characterised, except for some programmes in the social area, by dispersion and discontinuity. This is due to the two dominant characteristics of the country's government planning:

- (a) the emphasis on the process of formulating plans and developing programmes and projects; and
- (b) high negligence in the stages of monitoring and evaluating processes, results and impacts.

The above referred IDB / IPEA report presents a detailed mapping of discontinuous and diffuse monitoring and evaluation initiatives at five levels of planning.

The discontinuity also occurs in the records of data on evaluation. A brief analysis in the Activity Reports of IPEA (we have consulted the Activities Reports on IPEA's website from 2011 to 2019, see BRAZIL, IPEA 2020), points out that in 2019, 15 evaluations were carried out under the CMAP, 18 evaluations of the Social Development Goals indicators and 110 evaluations and proposals of the PPA. In 2018, the report did not present a division of the type of evaluation, and there were 72 evaluations in total. In 2017, at (the time when the CMAP was being discussed), 23 evaluations were recorded mentioning the SIAPRE (as described above SIAPRE was the first designed system by GTAG working group, which did not succeed). In previous administrations, there are no records of specific evaluations, and they present only numbers of publications in general.

What such data seems to indicate is that IPEA has been consolidating as an institutional reference within the federal scope of policy evaluations

for the last years. It is noticeable that the entity has a management data discontinuity, as its reports clearly show this. IPEA's Activity Reports can be used as a source to monitor the evaluations at the federal level and to check if such discontinuity will persist.

Besides the national policy at the federal level, states and municipalities also organise their evaluation systems. A study that examines this aspect is necessary, which during the interviews for this work revealed to be quite advanced in certain contexts. Some examples: the state of Espírito Santo, previously mentioned, with CLEAR's LAB advisory for monitoring, evaluation development and training (a project that overcame the recurring policy discontinuity due to the changes of the government); the government of Bahia has a solid tradition in the area of policy monitoring and evaluation and its team is part of the board of the Brazilian Network for Monitoring and Evaluation. The states of Ceará, São Paulo and the municipality of Belo Horizonte have also valuable initiatives in the development of education policies with evidence (FECOP Ceará, São Paulo Evidence Office and Evaluation Sector at the Education Department of Belo Horizonte respectively). The context of the institutionalisation of evaluation is permeated by the continuing political dispute. We have different perspectives on the design of public policies, including the debate on what should or should not be the responsibility of the state and how it should be guaranteeing rights and providing services to the population. The goals of a monitoring and evaluation system are also permeated by this structural dispute about the role of the state. Monitoring and evaluation can be carried out by: (i) **external** agents (or agency inside the government): a function closer to public resources management from a more auditory perspective conducted by The Federal or States Court of Accounts, (ii) **internal and external** agents (external agents such as the civil society, internal agents such as other government agencies or the actual program team). It can focus on improving public management in several phases: (a) from planning with data collection for **diagnosis** (or ex-ante evaluation) and policy design and (b) **evaluating processes**: monitoring and evaluating program's implementation for improvement. Another perspective, more in line with the debate proposed in this article, is (iii) **assessing impact**: the evaluation at the end of the policy or programme cycle, aiming at analysing and improving results. These initiatives are still very widespread in the public policy environment, it is not possible to identify a body that centralises the studies that have already been carried out, nor to account for whether more

internal or external studies are made and their nature: impact or process, for example. It is possible to note many studies conducted through different institutional arrangements:

- External Impact Evaluation: Ex post evaluation, Demand from international agencies such as the World Bank, IDB and other international agencies. Some examples: Impact Assessment of the *Bolsa Família*, Impact Assessment of the HIV/AIDS Programme,
- Internal evaluation, impact: Ex post evaluation, Demand from policy executing ministries, with technical support from academia/consultants: Analysis of Effectiveness Water for All Programme (partnership between Ministry of National Integration and Getulio Vargas Foundation Public Policy Analysis Board [DAPP FGV]).
- Internal evaluation, process: Evaluation of implementation conducted by the design of the policy itself, with technical support from other government agencies: Criança Feliz Programme, evaluation was conducted by IPEA.
- External evaluation, impact: Ex ante evaluation conducted by researchers independently, within universities and research centres. The *Brazilian Institute of Social and Economic Analyses (IBASE)* publishes booklet on the impact of port construction on the *Tapajós River*.<sup>20</sup>

The multiplicity of these arrangements shows how impact and result evaluation is still subject to the political agenda. If it is a very visible policy for society, the debate ends up expanding to more evaluations are carried out, contracted, et cetera. If a policy is not as visible, the demand becomes more focused on the technical environment and the possibility that the information extrapolates that environment is little. It's particularly interesting to think about the reflection carried out by some Brazilian researchers interviewed, about the need to expand the methodological range. Paulo Jannuzzi et al. (2018) address the term systemic evaluation, thinking about the evaluation dimensions proposed by the Organisation for Economic Co-operation and Development (OECD) and the need to

<sup>20</sup> <https://ibase.br/pt/noticias/ibase-lanca-cartilha-sobre-portos-no-rio-tapajos/>. Accessed in May 2020.

go beyond an analysis of impact (often referred to impact evaluations in the econometric formats, considered as the golden standard), as was the case of the Analysis of Effectiveness of Water for All Programme.<sup>21</sup>

The convergence of this knowledge and learning was and still is an important part of the agenda of many specialists in the field.

## USE OF EVALUATION

The use of evaluation in Brazil is still underexplored and under-researched. The evaluations prepared by the mentioned actors—as the academic world, the IPEA, the ministries, the CMAP and the evaluations funded by international bodies—can attract some visibility of the media depending on the popularity of the evaluated policy. The issue raised by many of the interviewees is that evaluation in Brazil has more of a diagnosis role, rather than being used to improve or correct public policy (one of the reasons been the discontinuity of programs).

There is no body in the country that gathers data on the evaluations conducted by the various sectors, nor does it indicate the use of the findings in the policies. The Court of Auditors of the Union has conducted the Report of Government Policy and Programme Monitoring (RePP) in recent years (2017, 2018 and 2019) in compliance with the 2019 Budget Guidelines Law (the law that complements the PPA of 2016–2019).

This report is a diagnostic evaluation of government programmes and actions. It points out the state of planning with regard to the monitoring and evaluation of actions—with the result of 59% of failures in monitoring and evaluation (lack of a monitoring and evaluation model and insufficient and low-quality data). The report also points out that the legislature can contribute to this improvement, making clear instruments for institutionalising these processes.

Another point is the lack of transparent information about current public policies, including objectives, indicators, goals and physical financial performance. There is a lack of result and impact indicators for part of the policies, lack of information to support monitoring and evaluation. In addition, the mechanisms used to promote communication and accountability do not ensure the necessary transparency.

<sup>21</sup> [https://www.researchgate.net/publication/336567872\\_Analise\\_da\\_efetividade\\_do\\_Agua\\_para\\_Todos\\_avaliacao\\_de\\_merito\\_do\\_programa\\_quanto\\_a\\_eficacia\\_a\\_eficiencia\\_e\\_a\\_sustentabilidade](https://www.researchgate.net/publication/336567872_Analise_da_efetividade_do_Agua_para_Todos_avaliacao_de_merito_do_programa_quanto_a_eficacia_a_eficiencia_e_a_sustentabilidade). Accessed in June 2020.

Although the Access to Information Law 12,527/2011 has been in force since 2011, it is still not possible to identify information on the ‘implementation, monitoring and results of programmes, projects and actions of public agencies and entities, as well as proposed goals and indicators’ in a complete and structured manner (BRAZIL, The Federal Court of Accounts 2019).

Finally, the report points out that it has not been identified in the federal government’s list of structuring systems, a system that would cover the entire cycle of public policies.

The report brings detail on the sectors with the best and worst performance, 58% (10 out of 17) policies analysed presented unsatisfactory results (3 out of 7 social policies, 2 out of 4 infrastructure policies, 3 out of 3 environmental policies, 3 out of 3 in science and technology). The numbers in the sample are not very significant, but they do shed light on the sector that is generally the most focused of evaluations (social sector), indicating what many interviewed specialists also approached informally: environment and S&T as one of the least evaluated sectors (BRAZIL, The Federal Court of Accounts 2019).

The proposal for this chapter is to briefly explore two cases of public policies evaluated in the country to understand certain aspects that enhance the performance of sector evaluation in the country. The democratic opening of the government, participation mechanisms, investments in research, interest from the media and visibility for international bodies are some of the contextual factors that enable the enhancement of the use of evaluations in the country.

Far from being a definitive debate, we believe that it is important to address the importance of involving multiple actors and how this can lead to changes and transformations of the public policies.

### **Bolsa Família Case**

The *Bolsa Família* (Family Grant Programme) became the main social programme in Brazil, the policy based on the transfer of income to the population in extreme poverty was instituted in 2003 integrating the social policy agenda that had been consolidating since 2001.<sup>22</sup>

<sup>22</sup> In 2001 the Bolsa-Escola programme was inaugurated and expanded by the PDSCE. The *Bolsa Família*, is therefore part of PDSCE, the central strategy of the Lula-Dilma governments.



Researcher Paulo Jannuzzi, in a recent work (Jannuzzi, 2020), makes an in-depth analysis of how the statistical system of Brazil benefited from this policy for its own development and improvement. According to him, the ‘political-statistical-policy agenda’ cycle was the focus of the period, since the diagnosis for implementing the policy was all based on data collected by IBGE, which also incorporated questions into the questionnaires or carried out new statistical surveys to supply information needs and evidence from the programme. It is important to emphasise that the Social Development and Fight Against Hunger Programme (PDSCF) were executed together with the strategic management centre of the government, a true transversal policy, involving several areas, ministries and agencies (the distribution of resources by the *Caixa Econômica* Federal and the Data Registry at the National Institute of Social Security). The policy’s relevance and centrality within the government brought the media attention and made it cover of the main newspapers, raising questions among the public opinion. The country faced a democratic dispute between the media, specialists in the sector, academia and the government itself of what could be analysed as the real impact of the *Bolsa Família* for the country. While common sense used pejorative expressions claiming that the programme stimulated laziness and even increased birth rates, the organised civil society used the data available through the statistical system to delve deeper into impact analysis, and the government itself also elaborated its studies through SAGI, a body highly respected at the time for its technicality and data collection system.

SAGI was responsible for evaluating and monitoring the social development policies and programmes of the MDS, which is an innovation of the Brazilian public administration, since until then there had not been a secretariat in any ministry with this exclusive purpose; above all, a unit located horizontally in relation to the finalist secretariats, and not vertically, as usually happens with evaluation and monitoring units. (Vaitsman et al., 2006, p. 15)

In addition to the continuous monitoring of the policy, many impact assessments were conducted, the largest of them with World Bank funding.<sup>23</sup>

<sup>23</sup> And here, there is a whole debate as to whether the methodology proposed by the study was the most appropriate to evaluate the programme.

What is most important at this period for our analysis goes beyond the conclusion that it was a period that provided much learning for the consolidation of public policy evaluation in Brazil. It's relevant to observe that although the evaluation of public policies can be done by a body centralised in the government, this process must be alive within the whole society.

Brazil lived the experience that the consolidation of a robust public policy naturally creates the need for accountability (whether through the media, civil society groups or electors)—in other words, people start questioning: is the policy effort effective? And from a democracy perspective, this debate must involve the whole society itself. The democratic openness and the possibility of weaving analyses by the media, civil society, academia, independent evaluators, enrich the analyses and advance the development of knowledge and learning capacities at a country level.

A government that avoids evaluation is weakening democracy. The maturation of a democracy depends on the capacity of a whole society to analyse its own data. As we have already said, this process also involves strong political elements and power arrangements, the evidence produced by evaluation is just another resource to guide decision-making or learning. In order to understand the complexity of a public policy context it is required to also assume that the such other factors will be present, but it is also necessary not to give up an evaluative analysis.

During the interviews for this work, important actors of this period highlighted the importance of data access and social participation to increase use of studies by the government. A particularly interesting case was reported by a member of a civil society research agency, saying that after the publication of one of his analyses by the press, he hears that the president (Lula, at the time) had called an emergency meeting to analyse the results of the research with his team. This is an information obtained informally by the researcher during the period of his research. There are still no mechanisms to verify the regularity in which the evidence is used for decision-making. The work carried out by evaluation agencies and training agents is more directed at creating a culture of evidence-based decision-making—which is still known to be under development by all actors—than establishing mechanisms to control the leaderships practices.

It is interesting to think that even though the use of evaluation may not be documented, the democratic openness and social participation can guarantee this process, as illustrated by the example. The institutionalisation of evaluation in the normative sense is therefore one of the ways

to homogenise this process, but the broad understanding of the use and meaning of policy evaluation is more extensive and involves the whole of society.

### *PMAQ—SUS Case*

This programme is particularly interesting as a case of development of the evaluation perspective on health. The programme uses the evaluation as a tool to enhance the quality of and widen access to health, involving the health teams in such process. This movement is very specific of a sector that provides services to the society, and its success is the guarantee of quality of such service. Therefore, health already has a connection to the evaluation thinking due to the nature of its activity. The understanding that this rationale must be expanded as to include the entire technical team in the evaluation process is also an interesting regard, which involves the teams in the ongoing evaluation process.

Due to its size and high cost, many criticisms were made to the programme, in the sense that it could induce the manager to dissemble to obtain funds (the transfer of health funds were connected to the results brought by the Programme), which increases bureaucracy and weakens the autonomy of the health professional that serves the users, for instance.

The need for an evaluation of the evaluation—that is, a meta-evaluation—arises from such questionings. The discussion about the meta-evaluation has been addressed by the specialists, who assert that there are lots of data to analyse and to understand the validity of the programme. The academic world has analysed such data and produced remarkably interesting materials on the topic.<sup>24</sup>

Besides providing input (data) for research on health evaluation, the programme itself is a learning opportunity for the manager. After three execution cycles, there is a nationally structured programme that may provide to the public manager knowledge on his health network and the necessary budget adaptation for such. During some interviews with specialists in the programme, reports were obtained on the use of the data

<sup>24</sup> See also: Lopes et al., 2015; Melo et al., 2016.

to enhance health in municipalities as *Ribeirão Preto*, which allocated the funds based on the analysis of the PMAQ data.<sup>25</sup>

The researcher Oswaldo Tanaka and Edson Tamaki, in their paper *O papel da avaliação na tomada de decisão nos serviços de saúde* (The role of the evaluation in the health service decision-making), talks about what is the actual meaning of the use of evaluation according to him, providing interesting arguments that show how health evaluation thinking has developed in the country:

The evaluation proceeding must seek that the decisions consider the health needs of the population, the goals defined by the services and involve the interested parties, as to make the implementation of the decisions feasible. (Tanaka & Tamaki, 2012, p. 821)

The above mentioned researchers have indicated that the ‘timing’ between performance of the evaluation and the possibility of implementation of its recommendations is critic, which requires funds and time, not always available for public management. For events that require immediate decision-making, only the accumulation of previous knowledge, arising from the practice of recurring evaluations, could contribute. An alternative approach is the concept of Executive Evaluation, addressed by the *Guia Avaliação de Políticas Públicas* (BRAZIL, Casa Civil et al. 2018a, p. 56).

The approach outlined by Tanaka and Tamaki (2012) and also by Matida and Camacho (2004) is that the evaluation has the purpose of supporting the decision-making and, so that can happen, applicability is part of the concept. Hence, the evaluator becomes a political actor that makes decisions of proceedings and choices of parameters so that the knowledge produced can reach the goal, supporting the decision-making. They mention some principles that must guide health evaluation, but that can surely be expanded for all evaluations that have as their final goal the decision-making by political actors.<sup>26</sup>

What happens when the evaluation is to be used for knowledge production, but without a clear goal on how its findings will be used?

<sup>25</sup> For more information on the programme, some authors such as Rogério Silva, Marco Akerman, and Oswaldo Tanaka are essential.

<sup>26</sup> They are: usefulness, opportunity, feasibility, reliability, objectivity, directionality (Tanaka & Tamaki 2012).

Tanaka, in an interview for this article, states that in this case what is being carried out is a **diagnosis** (some authors define it as an ex ante evaluation), but not an evaluation carried out to make decisions about the policy (already implemented). Every evaluation should be guided by an evaluative question: why do I need to conduct this study? What actions can be taken based on this knowledge? Interesting points to understand the development of evaluative thinking in Brazil, from the perspective of specialists with practice and experience in the national context.

### SOCIETAL DISSEMINATION

The private foundations are a relevant actor in this context, conducting projects that establish a connection with the public sector, states and municipalities, in many cases with the role of training such actors in the development of a policy in an effective manner. In this context, the goals are important to the execution of the projects implementation and the evaluation of the programmes was established transversally to the institutional activities. Itaú Social Foundation, *Roberto Marinho* Foundation, Laudes Foundation, *Maria Cecília Souto Vidigal* Foundation in the early childhood area are some examples of institutional work with a similar strategy and strengthening in the evaluation area. These are particularly interesting and complex examples of organisations specialised in management, with lessons regarding the monitoring of results that could be shared with other sectors. Based on interviews with certain advisors who are specialised in this sector, it was noted that the context develops independently and that some discussions have been had on that regard—such as on the meaning of the concept of *impact*, the inclusion of new methodologies and evaluation approaches.

In addition, the foundations develop work with municipalities focusing on the development of an evaluation culture with managers, as demonstrated by the case of the Lemann Foundation with *Formar* programme, which for almost five years has organised training with the education departments on several topics, including evaluation.

In the next paragraphs, we will analyse the practice of such organisations based on the mapping prepared by the Group of Institutes Foundations and Companies (GIFE).

### *Civil Society Demand for Evaluation*

Although some government monitoring and evaluation initiatives are being strengthened by civil society action, evidence indicates that civil society has had more of a role in training and supporting the public sector than in demanding evaluations through bureaucratic mechanisms of participation, for example. There is an understanding that it is necessary to create an evaluation culture and evaluation capacities in the policy implementer and not only normative instruments for them to occur. It's rare to see examples like the following case:

An interesting case of demand for evaluation by civil society is The Early Childhood Brazilian Network which develops a work of formation of the municipalities for the construction of early childhood policies with emphasis on planning and monitoring. The Network is composed by more than 200 civil society organisations working for the cause of early childhood in Brazil—from grassroot organisations to large organisations with their own funding. The Network plays an important advocacy role in the elaboration of Municipal Plans for Early Childhood, training public managers to elaborate and monitor the Plan. In addition, there is a strong relationship with the legislative sector, through the Parliamentary Front for Early Childhood with a consistent demand for increasing monitoring and evaluation practices in social programmes dedicated to early childhood. The evaluation of the Criança Feliz programme (see BRAZIL, Ministry of Citizenship, 2019, p. 21), for example, was closely followed by the Network.

There are also plenty more similar initiatives that, as we mentioned in the introduction, need to be analysed and mapped. The exchange of best practices could be more encouraged at the federal level, by using and exploring what already happens in partnerships between states, municipalities and the private foundations.

### *Institutionalised Use of Evaluations by Civil Society*

GIFE is an important institution for social investments in the country. It is the association of social investors in Brazil (institutes, foundations and companies) that has 160 members who invest, in total, approximately BRL 2.9 billion per year in the social sector, operating their projects or implementing third-parties' projects. For a better understanding of the social investment scenario, the association carries out a biannual census

with its members. In the 2018 civil society census, there were some questions related to the practice of the evaluation and considering 133 organisations that were answering 45% asserted that they had funds allocated for evaluation (mostly foundations and institutes 54% and other private social investment organisations such as companies).

Among the obstacles to carry out evaluations, it's possible to point out the high cost to develop a good evaluation and the belief that the impacts generated by the project are difficult to define and measure. Lack of time of the team was also repeatedly mentioned.

With respect to evaluation policies practices, 38% of the organisations affirmed that they have practices of disclosure for the external audience. A quite significant number, when considering the total of organisations that perform evaluations.

With regard to the value that the evaluation adds to the organisations of civil society, 67% affirmed that the most important goal of the evaluation is to identify the contributions of the project/programme. 44% affirm that evaluating all projects is an institutional policy, while 49% affirmed that they only evaluate the programmes with scale potential.

From the organisations that evaluate their programmes, 74.4% said that they perform evaluations with their internal team during the execution of the programme, and 44% engage external professionals.

Regarding the evaluation practices: 52% carry out a Theory of Change at some phase of the programme, 70% mention indicator matrix, but there is still no question about methodology.

20% of the organisations say that they do not evaluate their programmes, which is a number that represents how the sector has developed in the area and that there is still room to grow. There is still no question about the frequency with which evaluations are conducted within organisations, but the questions are more focused in terms of the existence of evaluation in the programmes implemented than in the period in which they are carried out.

### *Civic Participation*

Although the 1988 Constitution provided for several mechanisms of civil participation in the processes of elaboration and design of public

policies,<sup>27</sup> the RePP, prepared by the TCU, points backwards in this direction:

62. Decree 8, 243/2014, which established the National Policy on Social Participation (PNPS) and the National System for Social Participation (SNPS) was repealed in 2019 by Decree 9,759, which extinguished and established guidelines, rules and limitations for collegiate administration.

63. In spite of the discretion and legitimacy of the parties to address the issue, the fact is that the absence of a regulation on how to involve stakeholders generates gaps in the process of formulation, implementation and evaluation of public policies. (Items 62 and 63 of RePP, BRAZIL, The Federal Court of Accounts 2019)

So that the evaluation, as we are addressing in this paper, becomes part of such arsenal, the actors must be involved in its execution and its results must allow such actors to take a position in judgments that may translate into action. The reflection mentioned above was particularly important when talking about the use of evaluation in Brazil. Without a condition on its performance, the evaluation gets lost, the knowledge does not lead to action, and funds are wasted. Decision-making has always happened, with or without evaluation, taking into account strategic, political, economic, circumstantial (opportunities) issues. Evaluation must also be included as part of this decision-making arsenal in political leaderships.<sup>28</sup>

Despite the absence of mechanisms and practices for civil participation in evaluations, there is a very strong theoretical trend of social participation (mentioned by most interviewees) in health evaluation, there are some interesting experiences underway in the evaluations conducted by the Oswaldo Cruz Foundation in the Laboratory for the Evaluation of Escola Nacional de Saúde Pública Sergio Arouca (ENSP) Regional Endemical Situations and by the University of São Paulo (USP) in the Public Health Department.

<sup>27</sup> As participatory planning, through the cooperation of representative associations in municipal planning, as a precept to be observed by municipalities (Brazilian Supreme Federal Court (1988[1994]), Constitution of the Federative Republic of Brazil, Art. 29, XII); democratic management of public education in the area of education (Brazilian Supreme Federal Court (1988[1994]), Constitution of the Federative Republic of Brazil, Art. 206, VI); administrative management of Social Security, with the quadripartite participation of governments, workers, entrepreneurs and retirees (Brazilian Supreme Federal Court (1988[1994]), Constitution of the Federative Republic of Brazil, Art. 114, VI).

<sup>28</sup> This approach is also very well addressed by Goldman, I., & Pabari, M. 2020, p. 2.



Another great example of the participatory vision within evaluation is the study prepared by Thomaz Chianca and Claudius Ceccon 2017, two Brazilian researchers together with Michael Patton on the perspective of Paulo Freire in evaluative thinking. The authors bring the approach of liberation pedagogy as part of participatory evaluative thinking:

Even though Freire's most prominent ideas were developed before evaluation was considered as an established field, his approach of dialectic inquiry, critical reflection and taking action has influenced and still influences many leading scholars and practitioners within our (trans) discipline. While Freire did not use the phrase "evaluative thinking," his process of critical engagement with local people to analyse and understand their situation and take action based on their reflections manifests core elements of what today would be considered embedded evaluative thinking. In political reforms around the world, the retrospective evaluation of what happened in Guinea Bissau reminds us of the importance of evaluative thinking and reflective dialogue—and the fragility of both in the face of political polarisation and changes in leadership. (Chianca, 2018, p. 1)

### *Public Perception and Discussion of Evaluation Findings*

In Brazil, we have no concrete evidence that evaluations are perceived by the general public as an important tool for policy improvement. The instrument of evaluation best known by the population is the Census prepared by IBGE, very popular in Brazil (well approached by the media for generating temporary employment for young researchers).

The example below illustrates this mismatch between *policy results vs. media vs. public opinion approach*. This example illustrates how evaluation results are very rarely discussed in the media.

An example of the results dissemination of an evaluation is the study about the *Criança Feliz* programme (PCF), a programme of Social Development Ministry created in 2017 during Temer administration, right after the impeachment of the President Dilma Rousseff. The creation of the policy, focused on the early childhood, was a priority of the administration—it was led as a personal project of the minister at the time.<sup>29</sup> The evaluation of the programme became relevant since its design, and the

<sup>29</sup> The early childhood is a topic that has significant political force in Brazil, there is a diverse network of the organised civil society that massively pressures the governments for municipal policies and plans for the early childhood. The network was one of the

technical team sought to ensure institutional arrangements so that the evaluation was carried out even in the transition of administrations (Temer was in office for two years and four months). It is interesting to note that in this case the team prioritised some issues clarified by Tanaka, as the information speed and the time to assist in the decision-making. After the end of Temer administration, an evaluation of the programme, which had a budget of 350 million (surprisingly during times of fiscal adjustment), was published pointing out reservations to the programme, mainly on issues related to its implementation, as the communication with users and institutional arrangements:

Among the target audience, some families alleged that they would participate in the PCF for being afraid of losing the Bolsa Família (Family Grant Programme) if they refused to. This is a wrong idea, since the legislation that established the PCF (Decree No. 8,869/2016) only indicated priority to the families that benefitted from Bolsa Família (Family Grant Programme) to access the programme, not that this situation would be a condition for their participation.

In connection with the training events the interaction of the PCF professionals, beyond the mistrust of the visitors concerning the care for child development (CDC) and the Guide for Home Visit, for several times such visitors feel isolated within the programme, without connection with the coordinators and managers of their municipality or the teams from other places and, thus, the inclusion in such virtual areas would also work as to create a network of visitors and to form the shared identity. Many visitors miss access to higher instances and claim for more transparency in the expenditures and engaging manner.

Extracts of Notebook 34, Criança Feliz Implementation Evaluation. (BRAZIL, Ministry of Citizenship 2019, pp. 65f.).

Despite the interesting findings of the policy implementation evaluation, they were not discussed by the society in a broader sense: a search for

‘*Criança Feliz* Evaluation’ in the media<sup>30</sup> provided as results<sup>31</sup> with headlines that do not effectively present the results of the programme evaluation. All of them are positive news about the programme publicised by newspapers with low representation within the Brazilian large media. There are two cases, involving larger press vehicles, where the programme evaluation is addressed as a consequence of the fiscal adjustment of the government. In the second news, a gossip tone about the use of the image of the first lady to promote the programme. None of the news indicate the results of the implementation evaluation. The criticism to the programme does not reveal as relevant for the public opinion, nor for any change of its design, at least as informed to the society. The political force of the minister and of his programme have other meanings more aligned to power, not the effectiveness of the policy in people’s life. Before the begging of 2020 year, the plan was to increase the project budget to 800 million.

<sup>30</sup> For a better understanding on how the large media covered the campaigns, we used Google News as search engine. This is an engine that finds news published in newspapers, magazines, portals and other medias specialised in publishing news. It does not show commercial websites that use such tool for marketing, for example. We analysed the 100 main newspaper reports that addressed the program from 2017 to 2020.

<sup>31</sup> Examples of newspaper headlines about the program:

- “Criança Feliz undergoes globally unprecedented scientific evaluation” <https://diariodopoder.com.br/politica/programa-crianca-feliz-passa-por-avaliacao-cientifica-inedita-no-mundo>. Accessed in May 2020.
- “Criança Feliz wins the Wise Awards” <https://www.correiopovo.com.br/not%C3%ADcias/pol%C3%ADtica/programa-crian%C3%A7a-feliz-recebe-o-pr%C3%AAmio-wise-awards-1.381616>. Accessed in May 2020.
- “Government wants to attend to one million children by the end of the year with the Criança Feliz program” <https://www.camara.leg.br/noticias/581077-governo-quer-atender-um-milhao-de-criancas-ate-o-fim-do-ano-com-programa-para-primeira-infancia/>. Accessed in May 2020.
- “Lack of funds trigger public policy evaluation” <https://www1.folha.uol.com.br/mercado/2017/11/1932367-cofres-vazios-impulsionam-a-avaliacao-de-politicas-publicas.shtml>. Accessed in May 2020.
- “The use of Michelle’s image bothered Bolsonaro” <https://www.terra.com.br/noticias/brasil/politica/uso-da-imagem-de-michelle-por-osmar-terra-incomodava-bolsonaro,7026b7574428d9eb06f955a17e0112d8vvv4df0q.html>. Accessed in May 2020.

## PROFESSIONALISATION

One of the pillars of the training of the main Brazilian evaluators of these days was the private investment, and we highlight that W. K. Kellogg Foundation took the main evaluation advisors in Brazil to study in reference universities abroad. It is interesting to note this common training of specialists in several areas, such as health and social development, in the 1980s. In fact, the W. K. Kellogg Foundation report ‘Learning From The Field’ stands:

From 1942–71, WKKF’s programmes in Latin America and the Caribbean (LAC) were led from the foundation’s main office in Battle Creek, Michigan. This grantmaking emphasised the development of local LAC leaders through fellowships and leadership programmes from 1972–present, with enough trusted leaders in place, citizens of LAC countries led the programmes from the WKKF regional office, which was relocated several times:

- Rio de Janeiro, RJ - Brazil (1972–1985)
- São Paulo, SP - Brazil (1985–2009)
- Mexico City, DF - Mexico (2009–2018) (W. K. Kellogg Foundation, 2019, p. 2)

Another interesting indicator of the scenario of the professionals in Brazil is the events annually organised by CLEAR the gLOCAL Evaluation Week. A search on the official website of the event in 2019 and 2020 indicates 53 events organised by Brazilians. Some interesting data:

- Number of events: the number of events doubled in two years 18–35
- The number of participant organisations doubled from 12 to 23
- Considering those two years of events, there are 29 organisations in total: 9 academic entities, 8 advisories, 7 government, 4 civil society organisations and the Brazilian Network for Monitoring and Evaluation. This sample, although restricted, indicates that the evaluation market is dominated by academia and consulting firms, but with significant growth in the government sector.
- The number of governmental organisations presenting their works increased from 2 to 6. indicating increased government capacity to carry out evaluations
- And the presence of Civil Society Organisations presenting works in 2020 as Republica.org, *Arapiaú* Institute and *Ceará* Educative Territory Network.

About the content of the 53 events: institutional importance and evaluative culture (12), professional and youth training (7), exchange of concrete experiences (6), methodological innovation discussion—machine learning, behavioural economics and solidarity economy (4), impact questioning (4), tools (2), thematic debates: climate (2) education (2), quality (1), COVID (3), launches (2) and institutional bias (2).

- We note that the content of almost 25% of the events is still an introduction to evaluation, which shows the importance of the institutionalisation and the evaluation culture. The incentive for training and talking with young evaluators is also an expressive topic: almost 10% of the events. The exchanges of concrete experience are also significant, sharing learnings.
- An interesting data is that there are 4 events that address the myths of the impact evaluation in the econometric formats, many times considered as the gold standard. This information may indicate the emergence of new methodological formats based on the experience of the evaluators.
- There is still little discussion about tools and methodologies. And few sectorial discussions such as climate and education (they may be more concentrated in other thematic forums; the evaluation does not yet encompass these themes).
- The quality of the evaluation seems to be an emerging topic, and Brazilian Monitoring and Evaluation Network (RBMA) leads the discussion.
- Other uses of the events are the launching of publications and a more institutional bias.

Another important organisation is the RBMA, which monitors since 2014, the demand for evaluation in Brazil—aimed at the community of evaluation professionals (consultants and academia, mostly). The network is an established reference for professionals of the area and aims to share information and knowledge to strengthen the evaluation culture in Brazil. The Network is composed of more than 8,500 members with open access to the network, materials, job posting, webinars, etc. and 28 paying members who participate in the Assembly and administrative decisions. Analysing the RBMA database in two periods, from 2014 to 2016 and from 2017 to 2019, we note an increase of the demand for evaluations from international bodies, such as United Nations Development

Programme (UNDP), IDB, International Fund for Agricultural Development (IFAD) from 60 to 69% while the demand from state bodies, such as governmental departments and ministries, dropped from 17 to 9%. The demand from institutes and foundations changed slightly from 23 to 21%. Such follow-up is important for the knowledge and expansion of the sector's activities, as stated Marcia Joppert, RBMA founder and currently director:

Information management has indisputably proved to be a factor essential in networks, both in terms of seeking new links, creating greater possibilities to act in the dynamic context of the reality of what it does part and ensuring its maintenance and survival, as in creating internal control mechanisms that promote the best use of your resources. (Joppert et al., 2011, p. 8)

Despite the great potential of the civil society in claiming for evaluations, as indicated by the GIFE Census, the international organisms still have a central role in claiming for evaluation in the country.

The followings are some of the organisations that most claim for evaluations in the RBMA: IFAD, United Nations and United Nations International Children's Fund, UNDP, UN Women, Inter-American Development Bank and Laudes Foundation, besides John Snow advisory, in an average of 44 annual demands disclosed by the Network.

Jasmim Madueno (2019), a member of the RBMA and evaluator, analyses the professionalisation scenario in Brazil. According to her (i) there is a considerable offer of short courses available. However, longer courses are rare. The first master on evaluation of public policies was created in 2005 by the University of Ceará. Other two master programmes are available (Cesgranrio Foundation and National School of Public Administration), which is almost nothing taking into account the M&E field growth; (ii) the Brazilian Network for Monitoring and Evaluation has an important position on gathering and disseminating information about M&E; (iii) there is not a national definition about evaluation principles and guidelines, most of the evaluators orientate themselves by international standards; (iv) and (v) evaluation as a profession is not clear, neither are the competencies and backgrounds expected from those professionals. Recently there has been an update in the point (iii) with the launch of the 'Guidelines for Evaluation Practice in Brazil' (*'Diretrizes para Prática de Avaliações no Brasil'*). The Guideline has developed in a partnership

among The Brazilian Monitoring and Evaluation Network and Group of Institutes, Foundations and Enterprises—GIFE and is defined as ‘an instrument of dialogue for actors in the field of evaluation and an invitation for Brazilian society to expand its ownership and critical capacity on evaluations, intensifying its presence in the co-production of evaluation thoughts and practices in the country.’ The impact of this instrument on evaluation practice is yet to be observed.

Despite the limited offer of courses that have evaluation as their main theme, some research groups within universities in the areas of health, public administration, economics and education also explore the theme in a rich and comprehensive way, through occasional training, some examples, which are far from exhausting what exists in the country, not yet strictly mapped. Table 4.2 shows the main courses and trainings in the country.

**Table 4.2** M&E Courses in Brazil (Author’s development)

<i>Undergraduate</i>	<i>Masters</i>	<i>Extension/Technical Courses</i>
non-existent	Federal University of Ceará: Master in Public Policy Evaluation  Cesgranrio—Rio de Janeiro: Professional Master in Evaluation  ENAP: Professional Master in Evaluation and Monitoring of Public Policies	Cesgranrio Rio de Janeiro: Evaluation Management  <i>João Pinheiro</i> Foundation, Minas Gerais: Specialisation in Monitoring and Evaluation of Public Policies Centre for Learning on Evaluation and Results for Lusophone Africa and Brazil (CLEAR LAB), São Paulo: The Programme in Rural Monitoring and Evaluation, On Demand courses for public managers; UNIFAL, Minas Gerais: Evaluation of Public Policies and Social Projects Virtual Gov. School Distance Learning Course (ENAP): Impact Assessment of Social Programmes and Policies

## Summer Courses:

- Oswaldo Cruz Foundation (Fiocruz) Regional Endemic Situation Assessment Laboratory: Summer and Winter Course on Health Policy Monitoring and Evaluation
- São Paulo University—USP: Department of Public Health Practice, Faculty of Public Health Policy Management and Health

## Others:

- National Statistical School: Master's Degree in Population and Territory, approaching evaluation as a transversal theme

## Non Academic Trainings:

- *Itaú Social*: Economic Evaluation of Social Projects
- *Fundação Carlos Chagas*: Systemic Evaluation of Social Programmes

A place is still missing where such groups can meet and exchange experiences, mapping efforts have been carried out by the Brazilian Network for Monitoring and Evaluation there's still a fertile path to be traced. There is high demand for courses and learning in the area, in the coming years this scenario should be richer.

There is still no certification system for evaluators, nor the existence of an arbitration board or professorship.

There are some journals about evaluation in Brazil such as: *Revista Avaliação de Políticas Públicas* published by Universidade Federal do Ceará; *Revista Meta: Avaliação*, published by Cesgranrio; *Revista Brasileira de Monitoramento e Avaliação* with nine volumes published by the RBMA and the Ministry of Social Development (last volume released 2015); RBMA is republishing the *Revista Brasileira de Monitoramento e Avaliação*, now as *Revista Brasileira de Avaliação* (RBAVAL).

Some journals in the area of education that address evaluation: *Revista de Gestão e Avaliação Educacional* published by Federal University of Santa Maria; *Ensaio—Avaliação e Políticas Públicas em Educação* published by Cesgranrio, *Estudos em Avaliação Educacional* published by Carlos Chagas Foundation.



### *Evaluation Principles and Quality*

The Report on Government Policy and Programme Monitoring (RePP) (2019) does not mention any detailed analysis on the quality of the evaluations conducted. It mainly addresses failures in building monitoring and evaluation systems and public policy planning. The report's conclusions point in the direction that poorly designed policies also make an ex post evaluation difficult.

21. Analysing public policies, the object of this work, it is possible to identify multiple gaps in their formulation, gaps that range from the absence of guidelines and plans—that precisely define the expected results, objectives, goals and time frames of the policy—to errors in the inclusion of beneficiaries and weaknesses in the intervention logic.

33. In 2019 it was possible to identify the occurrence of gaps in the planning of 53% (9 out of 17) of the public policies analysed. These failures range from a lack of long-term vision and guidelines to weaknesses in the logic of intervention and policy execution without support for plans that define precisely the objectives, goals and expected results. (Excerpts from RePP report, items 21 and 33)

As mentioned above the Brazilian Monitoring and Evaluation Network in partnership with the Group of Institutes, Foundations and Enterprises—GIFE is inaugurating in 2020 a set of guidelines and standards for evaluation in Brazil. The initiative proposes to work with the evaluation field as a whole. Within a context, which as we have seen, the theme of evaluation is distributed among several sectors, such as health and education mainly. Associations such as Brazilian Association of Educational Assessment in education (recently extinct and absorbed by University of Campinas) and Brazilian Association of Collective Health in health—are constituted as sectoral focal points for evaluation thinking.

The culture of following evaluation standards is developing greatly. The main reference used most of the calls analysed on the RBMA platform, are the evaluation criteria of the OECD, Development Assistance Committee Network (Relevance, Coherence, Effectiveness, Efficiency, Impact, Sustainability) both by clients, organisations and evaluators.

## CONCLUSION

Brazil is a country with continental dimensions and very unequal realities. Its geographic and political complexity is also reflected in the analysis of the institutionalisation of policy evaluation in the country.

From a normative point of view, there is no legislation in the form of a law that regulates a national evaluation policy. Recently, the CMAP, an inter-ministerial arrangement that aims to monitor and evaluate public policies with the support of institutes such as the IPEA and other federal data collection bodies, was regulated by a decree. CMAP exits since 2017 and is a promising body in relation to the institutionalisation of evaluation in Brazil. It still not possible to observe results exclusively arising from CMAP considering that evaluations published by CMAP were already being conducted by other spheres of government before its conception. In addition to CMAP, it is important to highlight sectoral policies such as the PNE, which provides for evaluation as part of the policy cycle; the existence of SAGI, an evaluation office linked to the Ministry of Citizenship; and on a state level, *Espírito Santo* state, which just had a law approved institutionalising evaluation as a state policy. This latter example shows that there is a lot of potential for institutionalisation in the country to occur through state legislation.

The practice of evaluation is an important issue, since what actually happens in practical terms may be far from what is formulated by regulations or laws in force. It is possible to observe evaluation practices closer to a *diagnostic* function and less to a *learning process* about policies. It is still a frequent process used during the implementation of the policy plans, rather than used on other advanced phases. *Ex post* evaluations are recommended by international funding agencies, such as the *Bolsa Família* and the Ministry of Health's HIV-AIDS policies, which have invested in carrying out an impact evaluation by the university or the ministry itself.

How the results of the evaluations are being used by the population and decision-makers, is a difficult question to answer, as this matter is intimate connected to the political universe with a noted lack of information available. Given the interviews conducted with specialists, it was possible to observe that there was a particular period (2003–2014) where the social development policies were treated as priority matters under the government agenda, and such programmes started being evaluated by various actors in the public policy system. Universities, independent

researchers, research centres and civil society evaluations generated visibility for the data and public debate around the effectiveness of such evaluated policies. That context had a great influence on the visibility of a given policy and how public opinion still deals with it, as the case of *Bolsa Família* programme.

An organised civil society is a very strong actor when it comes to evaluation. Besides promoting, training and evaluating their own programmes, these organisations have the political independence to use the results of an evaluation to improve their programmes. Recent studies indicate that the ecosystem of social organisations in the country points to 133 organisations where 45% of organisations claim to have resources allocated for evaluation.

In relation to professionalisation, the country is still at the beginning of the expansion of the sector internally, offering some specific technical training for evaluation. The Brazilian Monitoring and Evaluation Network, CLEAR LAB and the GIFE are some of the training and associative initiatives for the professionals in this field. There are only three courses at the master's level that cover the whole subject.

Brazil has, therefore, a fertile environment for a growing institutionalisation of evaluation. There are important sector practices, qualified professionals involved in the subject and an increase in the development of effective institutional models that place evaluation as part of the policy planning and implementation cycle. Political will is however strongly associated with government cycles and trends and therefore it affects and influence how evaluation will evolve and become institutionalised.

### LIST OF ABBREVIATIONS

CGU	Comptroller-General of the Federal Union
CLEAR LAB	Centre for Learning on Evaluation and Results for Lusophone Africa and Brazil
CMAP	Committees for Monitoring and Evaluation of Federal Public Policies
CMAS	Monitoring and Evaluation of Union Subsidies
CMAG	Monitoring and Evaluation of Direct Expenditures
DAPP FGV	Getulio Vargas Foundation Public Policy Analysis Board
ENAP	National School of Public Administration
ENSP	Escola Nacional de Saúde Pública Sergio Arouca

ESAF	Government Budgeting Management School
FIOCRUZ	Oswaldo Cruz Foundation
FECOP	State Fund for Combating Poverty
GDP	Gross Domestic Product
GEF	Executive Management of FECOP
GIFE	Group of Institutes Foundations and Companies
GTAG	Interministerial Working Group for Monitoring Public Spending
IBGE	Brazilian Geographic and Statistical Institute
IDB	Inter-American Development Bank
IFAD	International Fund for Agricultural Development
INEP	National Institute for Educational Studies and Research Anísio Teixeira
IPEA	Institute of Applied Economic Research
IBASE	Brazilian Institute of Social and Economic Analyses
LDO	Budget Guidelines Law
MDS	Ministry of Social Development
MEC	Ministry of Education
NGO	Non-governmental Organisation
OECD	Organisation for Economic Co-operation and Development
PEC	Constitutional Amendment Proposal
PDSCF	Social Development and Fight Against Hunger Programme
PL	Law Proposal
PMAQ	The National Programme for Improvement of Basic Care Access and Quality
PNE	National Education Plan
PPA	Multiannual Plan
RBMA	Brazilian Monitoring and Evaluation Network
RePP	Report on Government Policy and Programme Monitoring
SAGI	Secretary for Evaluation and Information Management
SIAPRE	Public Policy Monitoring and Evaluation System and Public Expenditure Review
SIMAPP	Public Policy Monitoring and Evaluation System
SINAES	National System of Higher Education Evaluation
SUS	Unified Health System
TCE	State Court of Accounts

TCU	The Federal Court of Accounts
UNDP	United Nations Development Programme
USP	University of São Paulo

## REFERENCES

- Abranches, S. (1988). *Presidencialismo de coalizão: O dilema institucional brasileiro*, *Dados: Revista de Ciências Sociais*, 31(1), 3–55. IUPERJ.
- BRAZIL. [Constituição (1988)]. Constitution of the Federative Republic of Brazil. Brasília, DF: Presidência da República, [2016].
- BRAZIL, Casa Civil et al. (2018a). Avaliação de Políticas Públicas: Guia. Prático de Análise Ex Post. Instituto de Pesquisa Econômica Aplicada.
- BRAZIL, Casa Civil et al. (2018b). Avaliação de Políticas Públicas: Guia. Prático de Análise Ex Ante. Brasília: Instituto de Pesquisa Econômica Aplicada.
- BRAZIL, Ceará State Government. Decree N° 29,910, of September 29, 2009.
- BRAZIL, Decree No. 9834, of June 12, 2019. Instituto conselho de monitoramento e avaliação de políticas públicas.
- BRAZIL, Espírito Santo State Government. State Law N° 10,744/2017.
- BRAZIL, IPEA. (2002). Cooperação Técnica BID-Ipea: Fortalecimento da Função Avaliação nos Países da América do Sul. Relatório técnico / Pedro Luiz Barros Silva e Nilson do Rosário Costa. Instituto de Pesquisa Econômica Aplicada. [https://www.ipea.gov.br/portal/index.php?option=com\\_content&view=article&id=5437](https://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=5437). Accessed on 30 May 2020.
- BRAZIL, IPEA. (2020). Boletim de Análise Político-Institucional. Publicações-Boletim Institucional (ipea.gov.br). Accessed on May 2020.
- BRAZIL, Ministry of Citizenship. (2019). Notebook 34, Criança Feliz implementation evaluation. [https://aplicacoes.mds.gov.br/sagirmeps/ferramentas/docs/caderno-estudos-34\\_210x280\\_20200124-digital.pdf](https://aplicacoes.mds.gov.br/sagirmeps/ferramentas/docs/caderno-estudos-34_210x280_20200124-digital.pdf). Accessed on 30 May 2020.
- BRAZIL, Ministry of Economy. (1990). Law N° 8,029, 1990.
- BRAZIL, Ministry of Education. (2004). Law N° 10,861, of April 14, 2004.
- BRAZIL, Ministry of Education. (2007). Ordinance N° 660/201, 2007.
- BRAZIL, Ministry of Health. (2015). Ordinance N° 1,645 of October 2, 2015
- BRAZIL, Senado Federal. (2017). Project of Law N° 488/2017.
- BRAZIL, Social Development Ministry. Ordinance N° 2,227, of June 6, 2018.
- BRAZIL, The Federal Court of Accounts. (2019). Relatório de Fiscalizações em Políticas e Programas de Governo (RePP). Tribunal de Contas da União. [https://portal.tcu.gov.br/data/files/17/B0/92/77/E933071076A7C107E18818A8/Relatorio\\_politicas\\_programas\\_governo\\_2019.pdf](https://portal.tcu.gov.br/data/files/17/B0/92/77/E933071076A7C107E18818A8/Relatorio_politicas_programas_governo_2019.pdf). Accessed on 5 November 2020.

- Câmara dos Deputados. (1993). Art. 2º Resolution of Deputy Chamber Nº 48. Legislative Consultancy.
- Chianca, T. K., & Ceccon, C. (2017). Pedagogy in process applied to evaluation: Learning from Paulo Freire's work in Guinea-Bissau. *New Directions for Evaluation*, 155(1), 79–97.
- Chianca, T. K., Patton, M. Q., & Ceccon, C. (2018). Evaluative thinking in practice: Implications for evaluation from Paulo Freire's work. *Journal of Multidisciplinary Evaluation*, 4(30), 1–15.
- Grzybowski, C. (1991). Caminhos e descaminhos dos movimentos sociais no campo. Vozes.
- Goldman, I., & Pabari, M. (Eds.). (2020). Using evidence for policy and practice-Lessons from Africa. Taylor & Francis Group.
- Jannuzzi, P. (2020). Informação estatística e políticas públicas no Brasil: contribuições de pesquisas do IBGE para as políticas de desenvolvimento social e combate a fome (2004–2014), Universidade do Estado do Rio de Janeiro.
- Jannuzzi, P., Ruediger, M., & Meirelles, B. (2018). *Análise da efetividade do água para todos: Avaliação de mérito do programa quanto à eficácia, à eficiência e à sustentabilidade*. DAPP FGV.
- Joppert, M., Heriques, A., Pinho, J., Azevedo, J., Newman, J., Wenceslau, J., & Carvalho, S. (2011). Brazilian monitoring and evaluation network: Creation, Development and Perspectives. IOCE. [https://www.ioce.net/download/national/Brazil\\_BMEN\\_CaseStudy.pdf](https://www.ioce.net/download/national/Brazil_BMEN_CaseStudy.pdf). Accessed on 30 September 2020.
- Limongi, F. A. (2006). Democracia no Brasil. Presidencialismo, coalizão partidária e processo decisório. *Novos Estudos. CEBRAP*, 76, 17–41.
- Lopes E., Scherer, M., & Costa, A. (2015). O programa nacional de melhoria do acesso e da qualidade da atenção básica e a organização dos processos de trabalho. Tempus.
- Madueno, J. (2019). Current state of the institutionalization of evaluation in Brazil (Term paper), Saarland University.
- Matida, A. H., & Camacho, L. (2004). Pesquisa avaliativa e epidemiologia: Movimento e síntese no processo de avaliação de programas de saúde. *Cadernos De Saúde Pública*, 20(1), 37–47.
- Matuscelli, D. E. (2010). Ideologia do presidencialismo de coalisão. *Lutas Sociais, São Paulo*, 24, 60–69.
- Melo, L., Martiniano, C. S., Guimarães, J., Souza, M. B. D., & Rocha, P. D. M. (2016). Análises das diretrizes para o apoio institucional das gestões da Atenção Básica das capitais brasileiras. *Saúde debate*.
- Oliveira, P. (2019). O desafio da coordenação executiva pelo centro de governo: a experiência do Comitê de Monitoramento e Avaliação de Políticas Públicas (CMAP) de 2016 a 2017, Escola Nacional de Administração Pública.

- Silva, R., Furtado, J. P., Akerman, M., & Gasparini, M. (2015). Subsídios a meta: avaliações do PMAQ. In *Práticas de avaliação em saúde no Brasil: Diálogos*.
- Sodré, N. W. (1999). *História da imprensa no Brasil* (4th ed.). Mauad.
- Tanaka, O. Y., & Tamaki, E. (2012). O papel da avaliação para a tomada de decisão na gestão de serviços de saúde foi publicado na Revista Ciência & Saúde Coletiva, 17(4) Rio de Janeiro, editada pela Associação Brasileira de Pós-Graduação em Saúde Coletiva (ABRASCO).
- Torres, J. C. (1996). O encaminhamento político das reformas estruturais. *Lua Nova*, 37, CEDEC, pp. 57–76.
- Vaitsman, J., Rodrigues, R. W. S., & Paes-Sousa, R. O. (2006). Sistema de avaliação e monitoramento das políticas e programas sociais: a experiência do Ministério do Desenvolvimento Social e Combate à Fome. MDS. Unesco. Management of Social Transformations. Policy Papers, n. 17.
- W. K. Kellogg Foundation (2019). *Learnings and legacies from the field*. <https://wkkf.app.box.com/s/1kncpgnp0g8w6gplifh9dufj8x1dcrnm>. Accessed on 23 October 2020.



## Evaluation in Canada

*Benoît Gauthier, Robert E. Lahey, and Steve Jacob*

### COUNTRY OVERVIEW

Canada spreads over 9.2 million square kilometres of land and 0.8 million square kilometres of freshwater, for a total territory of almost 10 million square kilometres (Canadian Encyclopedia, 2015). This makes Canada the second largest national territory in the world after Russia. Canada is located immediately north of the United States. Its Northern neighbour, over the pole and the Arctic sea, is Russia.

---

B. Gauthier (✉)

Fellow of the Canadian Evaluation Society, Circum Network Inc., Cantley, QC, Canada

e-mail: [gauthier@circum.com](mailto:gauthier@circum.com)

R. E. Lahey

Fellow of the Canadian Evaluation Society, REL Solutions Inc., Ottawa, ON, Canada

e-mail: [RELahey@rogers.com](mailto:RELahey@rogers.com)

S. Jacob

Laval University, Quebec, QC, Canada

e-mail: [steve.jacob@pol.ulaval.ca](mailto:steve.jacob@pol.ulaval.ca)



As of October 2018, Canada's population was 37.2 million (Statistics Canada, 2018a). Recent population growth was mainly supported by international migration which accounts for 80% of the growth; the other 20% comes from the difference between births and deaths (so-called, natural increase). The share of natural increase has decreased steadily since 2012 and is expected to continue along this slope, mainly as a result of population ageing.

According to the World Monetary Fund, the Canadian economy is the 10th largest in the world, based on the nominal gross domestic product (International Monetary Fund, 2018). In 2018, 79% of Canadians worked in the service sector (Statistics Canada, 2018b). Within the service sector, 15% worked in wholesale and retail trade; 13% in health care; 8% in professional, scientific and technical services; and, 7% in accommodation and food services. Manufacturing accounted for 9% of all employment. Agriculture, forestry, fishing, mining, quarrying, oil and gas totalled 3% of employment. Public administration encompassed 5% of Canadian workers.

## INSTITUTIONAL STRUCTURES AND PROCESSES

The Dominion of Canada was founded as a federation in 1867. The British North America Act was enacted by the British Parliament (Canadian Encyclopedia, 2013). The Canadian head of state is the Governor General, that is, the representative of the British Queen; arguably, this is essentially a figurehead role leftover from the colonial background of the country (Docherty, 2014). Parliament is composed of the elected House of Commons and the non-elected Senate. The political party obtaining the majority of the seats at the first-past-the-post election is called upon to form the government. By tradition, the winning party leader becomes the Prime Minister and assembles a Cabinet which holds executive power and controls the public service.

The reality of the political dynamics is best explained by Savoie in this quote (2014, p. 135):

The executive has long held a dominant position in Westminster-style parliamentary governments. In formal constitutional terms, power is concentrated in the hands of the prime minister and Cabinet. Recent developments in Canada, however, suggest that the hand of the prime minister has been considerably strengthened. Indeed, when it comes to the political

power inherent in their office, it remains that Canadian prime ministers have no equals in the West. If anything, Canadian prime ministers have been able to strengthen their power still further at the expense of other political, policy, and administrative actors over the past several years.

The political structures of the ten provinces and three territories and the power dynamics within them are the same as what has been described for the federal government.

Most of the political and policy action in the federal public service takes place within some 70 departments and agencies, some 30 of which are considered large (Thomas, 2014). Departments are headed by a minister who reports to Parliament. Parliament votes operating and programme budgets annually. Given the stronghold of the governing party over decisions of Parliament in the context of majority government (which are the norm), in effect, the government is in charge of the allocation of resources as well as the implementation of programmes.

Canada's public service is qualified as 'professional' or 'Weberian' which means that it aims to support political decisions with the best possible analysis and to deliver on political decisions in a neutral manner. Nonetheless, criticism of the public service has been expressed towards an excessive power of the bureaucracy in the policy process, the threat that delegated administrative decision-making represents to individual freedom, and the overall problem that government performs poorly and wastes public money (Thomas, 2014). The professionalism and neutrality of the federal public service have been questioned in recent years (Chouinard & Milley, 2015).

Within this general institutional context, programme evaluation was born in the mid-seventies from a concern that the federal government had lost control over its finances (Jacob, 2006). The Auditor General report of 1976 described financial controls as dangerously insufficient and stated that it was indispensable to set up performance indicators to measure the effectiveness of programmes. In sync with the administrative culture, the introduction of the practice of evaluation was founded on the principle that department deputy heads are accountable for the performance of the programmes and the prudent use of public funds.

*Evaluation Regulations*

Evaluation in Canada has been strongly influenced by the federal public sector, in its evolution, structure and practice (Lahey, 2010; Lahey et al., 2018). Formalised systematic evaluation has existed in Canada for some 40 years, starting with the 1977 Policy on Evaluation, issued by the Canadian government. The Policy required every federal government department and agency to evaluate their programmes, following prescribed guidelines and standards. Though the Policy has changed four times over the past 40 years—in 1991, 2001, 2009 and 2016—it has served as the key document driving evaluation at the federal level.

With each change in the Policy, there has been an accompanying change in the mandate, focus and resourcing of evaluation (Lahey & Nielsen, 2013). This has had an impact on the amount and nature of evaluation that gets funded via the federal public sector. Table 5.1 below shows the level of commitment of the Canadian government to systematic evaluation, and how this has changed over the years. The data show that the demand for and investment in evaluation in the federal government is clearly affected by both political and economic factors such that there is an almost cyclical pattern in the spending on evaluation. Comparing annual federal expenditure on evaluation over 1983–2015 (constant 2015–2016 dollars) shows how shifts in policy have resulted in an ebb or flow for evaluation throughout the period. Also over the 2000s, the number of internal evaluators more than doubled (then tailed off somewhat). Both

**Table 5.1** Federal government commitment to evaluation (total expenditure on evaluation and total number of internal evaluators, selected years) (Lahey et al., 2018, pp. 49–51)

<i>Year</i>	<i>Expenditures (current \$M)</i>	<i>Expenditures (constant 2015–16 \$M)</i>	<i>Total no. of federal internal evaluators</i>
1983	22.7	49.2	n/a
1989	32.9	55.3	n/a
1991	26.1	39.9	n/a
2000	n/a	n/a	230
2001	n/a	n/a	285
2004	43	52.1	271
2009	62	68.7	488
2015	49.8	49.8	441

have an impact on the amount of evaluation that gets carried out, and the delivery model employed (internal vs. external).

But the model for how evaluation is institutionalised within the federal system has remained relatively stable. There are two key focal points for the delivery and use of monitoring and evaluation information: the Treasury Board of Canada Secretariat (TBS) that sets the rules and the individual government departments that measure the performance of their programmes and policies, making the Canadian model a poster child of the centralised policy-making and oversight plus decentralised implementation. Before exploring the current situation in some depth, let's look at the historical perspective. The significance of the role of the TBS as the central agency in charge of the practice of evaluation within the federal government has evolved over time.

- From the setting up of the federal evaluation system and initial policy in the mid-1970s to the coming to power of a new governing party in the mid-1980s, central influence on evaluation was high, with the Office of the Comptroller General acting as conductor, coach, focussing on training and development and setting standards.
- From the mid-1980s to the mid-1990s, under the Conservative Party government, expenditure reviews led to a scaling back of evaluation activity and a reduced central role for TBS with regard to evaluation.
- From the late 1990s to 2016, the central function of TBS was reinvigorated by two versions of the Policy (2001 and 2009), by the creation of the Centre of Excellence for Evaluation (CEE), by an influx of resources for capacity building and by mandatory requirements to conduct evaluations and to include a set of predefined questions in all evaluations.
- The policy centre for evaluation has gone through yet another transformation recently following the introduction in 2016 of the Policy on Results which led to a diminished central role for TBS and to the relaxation of mandatory requirements placed on evaluation. Among other things, there has been a move away from the comprehensive coverage requirement over a 5-year cycle which was found in the 2009 Policy; changes that would allow for evaluation scoping and timing to be based on needs of departments and not the mandatory core set of issues of the 2009 Policy; and, elimination of some requirements for evaluation for small departments and agencies. In

merging requirements of a modified evaluation policy and of performance measurement, the TBS dropped the CEE name and replaced it with a 'Results Division'. The role of this new central body is still unfolding. A key current priority is to address capacity weaknesses in performance measurement across federal departments.

The 2016 Policy (TBS, 2016a) and its accompanying standards and guidelines (TBS, 2016b) set the roles and expectations for evaluation. The central leadership role of TBS includes both oversight over the practice and use of evaluation across the federal system, as well as capacity building support for the federal evaluation system. Individual departments/agencies are given latitude in terms of their resourcing and use of evaluation, subject to the Policy requirements and standards. There are two types of standards for evaluation linked to the Policy—one that relate to the institutionalisation of evaluation within a department or agency; and, a second set dealing with the conduct, practice and use of individual evaluations. In that sense, the Canadian model is neither 'centralised' nor 'decentralised'; it is a combination of central rule-making and oversight with decentralised delegated implementation. Also important is that the oversight comes as checks and balances in a variety of centralised ways (TBS, the Management Accountability Framework, the Office of the Auditor General (OAG)) and decentralised ways (Departmental Evaluation Committees; neutral assessments of the evaluation function) (Lahey, 2010, 2011; Lahey et al., 2018).

One other federal policy that has had a significant impact on evaluation in Canada is the Policy on Transfer Payments, implemented in 2000 and updated in 2008 and 2012 (TBS, 2008, 2012). Issued under subsection 7(1) of the Financial Administration Act, it imposes a formal requirement to evaluate all programmes where the federal government funds, via grants and contributions, programmes that are implemented by other agencies, including other levels of government. Since transfer payments represent a large part of the federal government's spending and such programme funding typically covers a five-year period, a large number of programmes operating within the provinces are seeking evaluation services on a semi-regular basis. The Policy on Transfer Payments also stipulates a requirement for the formalised development of a performance framework and measurement strategy for each programme, intended to support the eventual evaluation that will assess issues of relevance and effectiveness of the particular programme.

While there has been a longstanding evaluation ‘system’ at the federal level, the institutionalisation of evaluation at other (sub-national) levels of government is quite uneven across Canada. Of the ten provinces and three territories that comprise Canada, only two (Newfoundland & Labrador and Quebec) have a formal policy on evaluation. Even in these cases, the policy, put in place in 2011 and 2014, respectively, has generated little actual systematic evaluation, due in part to budget cutbacks (Lahey et al., forthcoming).

That said, at the provincial level of government in Canada, there is a tendency for some form of evaluation to get carried out in ministries responsible for health, education and social or community services, though this is not necessarily the case across all provinces and territories. There is also a strong link with performance measurement and monitoring as a tool to track progress and performance (Lahey et al., forthcoming).

### *Evaluation Practice*

Canada is quite active in the field of evaluation at the federal level. Unpublished data from TBS (TBS, no date) about the federal government situation for fiscal year 2016–2017 indicate that 142 evaluations were completed and reported and that \$55.7 million (Canadian) were used by department and agency evaluation units (including salaries of 437 full-time equivalent employees which represent 78% of this amount as well as external contracts). These numbers do not include some 50 international development project evaluations.

As indicated above, the federal government has had five evaluation policies over four decades. Arguably, the 2009 policy has had the most impact of evaluation practice because it set “stricter requirements on evaluation”, which increased the demand (Lahey et al., 2018, p. 48). One of the major goals of this policy was to evaluate every major programme of the federal level (Lahey et al., 2018). To do so, the evaluation services had to be increased to respond to the higher demand (Lahey et al., 2018).

The 2016 Policy on Results relieved departments from many of the constraining requirements to their practice (Lahey et al., 2018) although the straitjacket remained for the evaluation of grants and contributions programmes under the Financial Administration Act. The relaxation of constraints was in part the consequence of the observations of an evaluation of the 2009 Policy on Evaluation (CEE, 2015).

The reach of evaluation activity has varied over the years, depending on the willingness of the federal government to invest in the evaluation sector. Generally, evaluation practice is more developed in “the field of education, health, labor markets, social policy, aid, industry policy, environmental policy, and research and development” at the federal level (Jacob et al., 2015, p. 14). At the provincial level, demand of evaluation is relatively low except for the provinces of Québec and Newfoundland & Labrador and outside the areas of education, health, and social services (Lahey et al., 2018). Resources are also seen as a major obstacle for the provinces. In provinces and municipalities, the demand for evaluation depends on the judgement of managers on the usefulness of conducting an evaluation. However, their limited knowledge in evaluation techniques is preventing an effective use of evaluation at these levels of governance (Greene, 2002).

A clearer distinction between the roles of internal and external evaluators would be beneficial according to a study done in 2000 by the Centre of Excellence for Evaluation. For 40 years, the federal government has relied upon both internal and private sector evaluators to evaluate its performance (Lahey et al., 2018). A majority of Canadian evaluations are done internally by agents who are members of the administration in charge of the implementation of the programme under evaluation (Jacob, 2006, p. 524) and who act as independent ‘critical friends’ (Fetterman, 2005) within the administration. There was a time when internal evaluators acted as ‘managers of external evaluations’, instead of carrying the evaluations themselves. Because the demand for internal evaluators has risen in the government sector, many young employees of private firms have moved to employment in the federal public sector (Lahey et al., 2018). This, coupled with reduced contracting budgets, has led to fewer delegations of evaluations to the private sector, to the contracting of components of evaluation (for example, surveys, focus groups, peer reviews) instead of entire studies, and to an increased level of substantive involvement of internal evaluators (Lahey et al., 2018). While there is no formal evidence on the real level of independence of the internal evaluation units, the perception is that they are granted latitude to do professional work. However, it is not uncommon that there be negotiation of recommendations flowing from evaluation findings between evaluators and programme managers, and even of the exact wording of some aspects of evaluation reports given that these reports will be

publicly available. Internal evaluators are not subject to formal professional requirements, but a competency framework for three levels of evaluator positions was issued in 2019 by the Treasury Board of Canada Secretariat. Also, in addition to central standards for evaluation and TBS oversight, individual hiring practices within departmental evaluation units build professional requirements into job descriptions and hiring practices. Every department has a senior Departmental Evaluation Committee that oversees the practice and use of evaluation. There is a requirement to carry out a ‘neutral assessment’ of the practice and use of evaluation on a regular basis (Lahey, 2010).

There has been a push towards performance management since the 1990s to respond to the higher demand for accountability and outcome management (Lahey et al., 2018). Several initiatives have been created over the years motivated by the desire to know “what works”, which led the evaluations to be “more [interested] about performance management” than evaluation itself (Lahey et al., 2018, p. 50). In Canada, audits also have an important role in performance evaluation (Jacob et al., 2015). Because of the movement of public management as well as the Results for Canadians approach from the early 2000s (TBS, 2000), evaluations are more and more focussed on outcomes and results and use indicators that are, elsewhere, implemented into national laws that make outcome and impact evaluation mandatory (Jacob et al., 2015); not the case in Canada however. Without replacing older practices in the field of evaluation, impact evaluation has made its way into the tools available to Canadian evaluators (Jacob et al., 2015). An indicator of the popularity of results evaluation is the fact that 94% of evaluations completed at the federal level since 2001 were oriented towards the success of the programme whereas 72% focussed on implementation (Jacob, 2006).

The 2009 Policy required addressing the issues of programme performance and relevance; this led to modifications in evaluation practice to ensure that all aspects of this new type of evaluation were covered (Lahey, 2010). Performance measurement (in Canadian lingo; ‘monitoring’ elsewhere in the world) was always part of management responsibilities within results-based management but it took on new life in the 2000s with the measurement of outputs as well as outcomes (Lahey, 2010). In addition, cost/benefit analyses have gained in popularity (Greene, 2002). In effect, process and impact/outcome evaluations are used in symbiosis, but accountability requirements increase the demand for performance and result evaluations.



The CEE was created within TBS with a mandate to assume a leadership role in the advancement of evaluation practices in the country. The CEE developed several initiatives in the early 2000s under the umbrella of a community development strategy, some of which survived, others not: competency profiles for federal evaluators; an Internship Programme for new evaluators; a training and development curriculum for in-career evaluators; a capability to monitor the health of evaluation government-wide. CEE also played a role in facilitating the bringing together of the Canadian Consortium of Universities for Evaluation Education, established in 2008 as part of a multi-organisational strategy to increase opportunities for current and prospective evaluation professionals to acquire the knowledge and skills required to become evaluation practitioners. The CEE was disbanded in 2016 following the absorption of evaluation within the Policy on Results.

There is little recent information on evaluation practice in the provinces (see Gauthier et al., 2004, 2009). In the province of Ontario, ministries have internal evaluators who “carry out a range of evaluation work” (Lahey et al., 2018, p. 55). However, when the subject of the evaluation is more politically sensible, it is common to ask external evaluators to conduct the work to prevent any misconceptions about the neutrality of the evaluation results.

Academics also take on evaluation work in Canada. Although most of them are more research-oriented and focus on the long-term, some actually do evaluations for smaller-scale organisations or provincial and municipal governments (Lahey et al., 2018).

### *Use of Evaluations*

Evaluation in the federal government can and does play a variety of roles, serving the needs of different users at different points in time. In broad terms, the uses of evaluation have typically gravitated between the support of programme or senior management—what might be considered the ‘learning’ role for evaluation—and their support of ‘accountability’ to senior or central agency authorities or programme funders. Over the 40 years of systematic evaluation experience in Canada, there has sometimes been a tension in finding the right balance between these two roles (Lahey & Nielsen, 2013).

For example, the 1980s and the 1990s saw the evaluation function tending to support programme management—through providing greater

knowledge about programme operations, supporting programme development and assisting in the development of performance indicators. But, too little examination of programme effectiveness was carried out, causing the evaluation function to be criticised by the Auditor General for not meeting its full mandate (Auditor General of Canada (OAG), 2009; Lahey, 2010).

The 2000s have seen a movement towards the use of evaluation increasingly for purposes of ‘accountability’ (i.e. are programmes delivering intended ‘results?’). This has been true in all sectors. In the federal public sector for instance, the 2009 Policy on Evaluation placed a focus on government-wide needs, forcing departmental evaluation units to focus on a stricter and more centrally oriented set of issues and, in the process, significantly lessen the ‘learning’ and management support role that had dominated the two previous decades. The introduction by government in 2006 of the Federal Accountability Act had set the tone for this focus on accountability in evaluation.

Another influence, the government’s emphasis on expenditure management and strategic reviews, starting in 2005–2006 and running in one form or another over several years, reinforced the aim of having 100% evaluation coverage of all major programmes, and thus served to support an increased need for evaluation services within the federal government sector (Dumaine, 2012; Lahey, 2015). The cycle formally requiring that all government direct programme spending be evaluated every five years proved to be a double-edged sword as it raised the profile of evaluation both within departments and centrally but it also forced the use of evaluation to principally serve central agency needs, sometimes/often neglecting department-level priorities (CEE, 2015).

The current 2016 Policy would seem to be attempting to establish a better balance between the two broad uses for evaluation, allowing as it does some flexibility for government departments in identifying priorities for evaluation and incorporating issues of concern to programme managers—while still noting the importance of central evaluation issues of relevance, effectiveness, and efficiency. The current policy still does note that the evaluation function in departments could be called upon by central authorities to “support resource alignment reviews” (TBS, 2016a).

Use of evaluation findings has been greatly aided by formal requirements within the federal policy on evaluation. Managers whose programme area was the subject of an evaluation must formally respond

to the recommendations of the evaluation with a ‘Management Response and Action Plan’ (MRAP) relating to each specific recommendation (TBS, 2016b). These are tabled with a very senior committee—chaired by the Deputy Minister or their delegate—and monitored over time to ensure their implementation. This high profile helps ensure follow-up to the actions being recommended by an evaluation.

The link of evaluation to Parliament generally comes in the form of reporting and in various ways (Lahey, 2010). At a macro-level, and on an annual basis as part of the Government’s Estimates exercise, Departmental Results Reports (DRR) are tabled in Parliament. Each department is required to prepare a DRR, based on, among other things, the results of programme evaluations, so as to inform Parliamentarians and Canadians of the actual performance and results achieved against plans by government organisations. Results of individual evaluations can and do get brought into discussion by relevant all-party Parliamentary Committees. Nonetheless, the national parliament and parliaments in provinces and territories do not play a significant role in the demand or the supply of evaluations.

On a more micro-level, programme spending/renewal proposals being submitted via Memoranda to Cabinet or Treasury Board submissions are required to be accompanied by the results of relevant evaluations to assist the decision-making. Unpublished data from the TBS for fiscal year 2016–2017 indicate that 149 of 421 (35%) funding submissions to TBS included evaluation findings and that 62 of 444 (14%) Memoranda to Cabinet also did.

More recently, with the introduction of a ‘Results and Delivery’ approach by the Liberal government elected in 2015 as a mechanism to track progress against government priorities, evaluation is starting to be called upon in some federal areas to assist the efforts in measuring high-level results linked to strategic priorities (Lahey, 2017).

But all is not well. Chouinard and Milley concluded “The paradox of use refers to the fact that, despite the discourse surrounding evidence-based practice and evidence-based decision-making, the actual use of evidence seems to be in decline, while the infrastructure surrounding evaluation (what we have referred to as the evaluation industry) seems to be on the rise” (2015, p. 14). Similarly, for a decade of Conservative Party ruling (2006–2015), the Canadian landscape of evidence-based policy-making was bleak (Thomas, 2014). By way of example, the Government of Canada used to have an Economic Council, a Science Council, a

National Roundtable on the Economy and the Environment, among others (Doern, 2007); those evidence-producing organisations and others were disbanded.

In the Not-for-Profit (NFP)/Voluntary sector in Canada, where evaluation services started to take on more prominence over the 2000s, the focus has largely been on ‘accountability’, as funders at every level, began demanding that social service organisations demonstrate effectiveness and impact of their programmes. A recent report on evaluation in the Ontario NFP sector concludes that 96% of Canadian charities evaluate their work in some way; organisations most commonly evaluate their outputs, outcomes and quality; when it takes place, impact evaluation is in addition to these concerns (Lasby, 2018, 2019). The Ontario NFP sector uses evaluation findings most commonly to report to Boards of directors and funders and to learn about programme objectives and outcomes (Lasby, 2018). The challenge for many NFP organisations though has been the lack of resources, skills or experience to reach this level of measurement (Lasby, 2019). While there is increasing recognition within some NFPs of the utility of evaluation as a ‘learning’ tool—whereby the organisation (and/or the funder) could learn and implement programme improvements, much akin to formative and developmental evaluation—limited budgets, rather than a lack of interest, would seem to minimise this role for evaluation in this sector to date (Lahey et al., 2018).

As noted earlier, much of the evaluation that gets carried out at the provincial level of government is associated with the accountability requirements of grants and contributions from the federal government. Beyond this though, there has been a tendency, though uneven across provinces, for evaluative efforts to be undertaken within ministries associated with health and education.

In summary, the main users of evaluation in Canada are senior managers and programme managers, both being focussed on instrumental use.

## SOCIETAL DISSEMINATION/ACCEPTANCE

### *Institutionalised Use of Evaluations by Civil Society*

While the general public (or ‘civil society’) in Canada would generally not initiate a formalised systematic evaluation, there are many formal and informal elements that bring evaluation into the public sphere.

One such element is stakeholder participation in the conduct of an evaluation. As standard methodology for evaluation in Canada, for example, it is expected that stakeholders who are in some way implicated with a programme, as a beneficiary or user of programme services, or implicated in programme delivery, would be consulted during the evaluation. The Standards adopted by the Canadian Evaluation Society (CES) speak to this (Yarborough et al., 2012, p. 23): “Evaluations should devote attention to the full range of individuals and groups invested in the programme and affected by its evaluation” and “Evaluations should be understandable and fair in addressing stakeholder needs and purposes”. CES standards on transparency also serve to reinforce the notion of “full public disclosure” for evaluation: “Evaluations should provide complete descriptions of findings, limitations, and conclusions to all stakeholders, unless doing so would violate legal and propriety obligations” (Yarborough et al., 2012, p. 139).

This is also formally built within the federal policy, where the Policy on Results specifies within its Evaluation Standards the requirement that evaluations “include sufficient and appropriate consultation with major stakeholders” (TBS, 2016b).

Ensuring transparency of evaluation is another element that works to bring evaluation closer to the general public in Canada. At the federal level, a specified directive associated with the government’s Policy on Results formally requires that all departmental evaluations must be posted on the department’s website, along with the programme response as reflected in the MRAP (TBS, 2016b). The policy went further than previous versions by requiring the posting of evaluation summaries, in addition to the full report—the intent being to support uptake by increasing user-friendliness of evaluation information.

The Access to Information Act, enacted by Parliament in 1985 (last updated in 2014), reinforces transparency by providing individuals the right of access to information under the control of a federal government institution. Publicly reporting evaluation results in Parliament is another mechanism that can serve to bring evaluation into the public sphere (see section ‘[Use of Evaluations](#)’ within ‘[Institutional Structures and Processes](#)’ for details).

Parliamentarians, indeed all Canadians, will also hear of the performance of government operations via annual audits of the Auditor General of Canada (AG) that get tabled in Parliament and become very high profile and public documents, widely reported in the media. These AG

reports are not effectiveness evaluations per se, but do support and enable an environment where objective analysis and constructive criticism, whether delivered via an AG report of government operations or a neutral evaluation of a government programme, is the expected norm. The AG's periodic system-wide audit and reporting to Parliament on evaluation implementation across the federal public sector serves to raise awareness of the importance and use of evaluation in the public sector.

One final note is to clarify the concept of 'civil society' and the use of this term in Canada. Internationally, particularly in the international development sphere, Non-Governmental Organisations (NGOs) are often referred to as representing 'civil society'. Comparable to NGOs in Canada would be organisations in the NFP sector. As noted earlier, this sector does initiate evaluation and evaluative work (though often constrained by resources), but at best, it would only be considered a component of what might be deemed to be reflective of 'civil society' in Canada. Actual use of evaluations by civil society in Canada is not formally documented but there is no significant manifestation of massive use of this nature. Anecdotally, one can find instances where evaluations may be used by professional groups or associations for symbolic use. While federal evaluation reports are freely available on government websites, they are not the subject of active promotion or of a communication strategy.

Use by associations, professional groups, programme beneficiaries and citizens is limited and tends to be for symbolic use.

### *Public Perception and Discussion of Evaluation Findings*

The word Evaluation is not commonly used in the public discourse. 'Reviews' and the more general 'research' are found more often in the media. In French, the word *Évaluation* is regularly used in the press—outside of the education sector where it refers mostly to student assessment. It mostly refers to the environmental impact assessments that are conducted from all major industrial projects. A search in the newspaper archives of *Le Devoir* (a highbrow liberal printed media) found 1,315 references to the word *Évaluation* in 2018. Our analysis of the first 50 references indicated that 21 referred to the assessment of ecological impacts followed by 7 references to financial assessments. At best 4 references could have been in relation to evaluation as conceive of it in this document.

A search in the Parliament of Canada Hansard for the word Evaluation returned a single reference for the whole year of 2018. Nonetheless, as stated elsewhere, evaluations are used as input into the reporting on government programmes and as sometimes reported in Parliamentary Committees.

The media are either not tuned into the evaluation culture or they lack the resources to use evaluation information. While they pay attention to reports produced by ‘watchdogs organisations’ such as the Auditor General of Canada or think tanks like the Conference Board of Canada, the media make basically no use of the federal evaluation reports even though they are made publicly available on department and agency websites—by requirement of the federal Policy (the Deputy Head is responsible for “Ensuring that evaluation reports and summaries, including complete management responses and actions plans, are released on web platforms as prescribed by the Treasury Board of Canada Secretariat”, TBS, 2016a, paragraph 4.3.18). Without empirical evidence to this, the situation could be explained by a perception that federal evaluation reports being produced internally they are not credible sources of analysis or criticism, or by the fact that evaluation reports are not accompanied by press releases that make it simple for media to report on them. This is in stark contrast to the level of communication, promotion and media reporting of the Auditor General reports which may be explained by the interest that the Auditor General has that its reports be very public (whereas government departments have no interest in public discussion of their evaluation reports).

Demands from civil society and private enterprises for evaluation are few and far between. There are no formal sources of information on this phenomenon but recollection from these authors could not produce a sizable number of instances where public decisions were criticised on the basis of lack of empirical evaluation evidence.

## PROFESSIONALISATION

### *Academic Study Courses and Further Training*

Training in evaluation exists in many forms in Canada. It is classified in two categories: academic training and professional development.

Universities offer training in evaluation at the undergraduate, master and doctoral levels. Such training can target students who are interested

in an initiation to evaluation or in advanced courses. Some specialised graduate programmes provide comprehensive training for future evaluators. For example, the Quebec *École nationale d'administration publique* (National School of Public Administration) offers a masters degree and a doctorate degree in evaluation; the University of Ottawa, Carleton University, the Université du Québec à Montréal and the University of Victoria offer graduate diplomas in evaluation; the University of Waterloo has a Master of Health Evaluation.

Everywhere in Canada, university programmes offer individual courses specialised in evaluation. The Consortium of Universities for Evaluation Education has put together a 77-page inventory of available courses (Hunter & McDavid, 2018). Some of the specialised training in evaluation techniques takes place in other disciplinary areas such as public health, criminology, public administration and psychology. For example, Laval University and the University of Toronto offer evaluation programmes in public health. Other universities offer programmes that are designed to study evaluation in different fields. For example, Ottawa University offers a graduate programme in interdisciplinary programme evaluation. This approach ensures evaluation training beyond students in public administration, into education, nursing, health and public health, social work, political sciences, economics, urban development, nutrition, management, criminology, and leisure, culture and tourism.

Several university professors make a career of teaching evaluation principles and practice; a few are located in programmes specific to evaluation while most are housed in disciplinary programmes. Academic researchers are active in the field of programme evaluation in almost all provinces and they directly contribute to evaluation training. In particular, three Canada Research Chairs (CRC) have been awarded in the area of evaluation and have created nodes of academic training: the CRC in Evaluation and Health Care System Improvement; the CRC in Programme and Policy Evaluation; and the CRC in Evaluating Public Actions Related to Young People and Vulnerable Populations. CRC holders are highly regarded academics with a history and a future of excellence in their field.

Professional development activities address new and seasoned evaluators who need specialised training over the course of their career. These training opportunities are provided by universities, private institutes, some government entities (like Statistics Canada or the Canada School of Public Service and some internal training by departments), the CES and its regional chapters and other like-minded professional associations. They



are numerous and diverse; they use a variety of approaches, from in-person delivery to online self-training. On its website, the CES maintains a non-exhaustive list of upcoming training opportunities.

The CES' own professional development includes basic training targeted at beginners called the 'Essential Skills Series in Evaluation'. This 4-day seminar outlines the elementary concepts of evaluation to prepare students to become professionals in evaluation. To accommodate evaluators who cannot attend in-person training (Canada is a large country), the CES e-Institute offers online courses developed by evaluators in English and French. The CES also considers its annual conference as a key professional development opportunity. The first conference took place in 1980. It is now attended by approximately 500 evaluators, most of whom deliver papers, posters, panels and debates.

Continuing education is a requirement for evaluators holding the CES professional designation of 'credentialed evaluator'. Over a three-year period, they must provide a report which demonstrates that they have accumulated at least 40 hours of professional learning. Evaluators are encouraged to participate in different evaluation activities so as to possess a diverse set of skills.

Professional development also takes a competitive form for Canadian students who participate in the 'CES/CESEF Student Evaluation Case Competition'. The CES and the Canadian Evaluation Society Educational Fund (CESEF) charity manage this competition which has been in place for more than two decades. In teams of three to five, participants have to analyse an evaluation case file and they are given five hours to return a professional-grade response. In the first round, teams work from their base location. Three round-one teams are selected for the final round which is held at the CES annual conference; students are subsidised for travel and accommodations. Teams are once again given five hours to study another case and to deliver a live presentation in front of judges and of an audience. The Canadian Journal of Programme Evaluation has devoted a special issue to the Student Case Competition (volume 18, Spring issue).

The CESEF has also created an award called 'Student Excellence Advancing Evaluation Knowledge Award', which is given in partnership with the CES. The creation of this Award seeks to encourage students' contributions to the domain of evaluation. To be eligible, students must be nominated by their teachers. The CESEF further supports evaluation

training with scholarships, travel grants, awards and educational opportunities. The goal of this non-profit organisation is to promote the practice of programme evaluation in Canada by assisting students who pursue studies in this field.

### *Profession/Discipline*

The Canadian Evaluation Society (CES) was officially created in 1981 (CES, 1981) although it held its first annual conference in 1980 (CES, 1980). It is the oldest national evaluation association still in existence. Starting small, it reached a membership of close to 2,000 individuals in the 2000's; as of February 2019, the CES has more than 1,800 paid-for members.

The CES' mission is to promote the development of evaluation theory and practice, to lead the professionalisation of evaluation, to build awareness of evaluation and to advocate for the use of quality evaluation.

The CES is based on twelve regional chapters (one per province plus one for the National Capital Region and one for the Yukon Territory) which are responsible for delivering most in-person activities (training, networking). The CES Board of Directors comprises one representative from each chapter plus a president and a vice-president who are elected by members and a treasurer who is selected by the Board. The Board is organised in a small number of standing committees and supported by various working groups: the Fellowship, the Credentialing Board, the Conference Committee, the Diversity Working Group, the Sustainability Working Group, the International Working Group and the Sponsorship Working Group.

The CES is active on a number of front including professional development, networking, knowledge development and advocacy. Its main products are an annual conference, a journal, a credentialing programme, webinars, an online training institute, in-person training, networking events, an active website and a mentoring initiative. The CES issues a weekly newsletter to its members.

Published under the auspices of the CES, the Canadian Journal of Programme Evaluation is a bilingual (English and French) open-source peer-reviewed publication that was first published in 1986. It seeks to promote the theory and practice of programme evaluation by publishing full-length articles and shorter practice notes. The journal has a particular interest in articles reporting original empirical research on evaluation. For

many years, it has produced three issues per year—two regular issues and one thematic issue. Each issue contains six to eight articles plus book reviews.

Canada was the first country where a professional designation was created for evaluators. The CES Credentialing Programme was launched in 2010 after substantial debates, controversies and deliberations in Canada and the United States. The programme planning and initiation benefited from a convergence of positive forces that eventually tipped the scale in favour of greater rigour in the recognition of the professionalisation of the evaluation trade (Love, 2015). The CES commissioned a think piece on the options it could consider to support the professionalisation of evaluation practice. The report recommended that the CES establish three levels of professional designation (member, credentialed evaluator, professional certified evaluator), each with progressively more demanding criteria (Halpern et al., 2015). The CES elected to focus on the credentialing option. Following a disciplined, participatory and sometimes difficult process, a special group of CES volunteers conceived the credentialing programme between 2007 and 2009 (Buchanan, 2015). One mammoth task was to develop an agreed-upon set of competencies that would be used to gauge the competence of applicants to the credential. The competency development process was based on existing literature and on a highly inclusive member participation process. This led to the establishment of 49 competencies organised in five domains of competence (Maicher & Frank, 2015). The competency scheme was approved in 2008 and revised 10 years later. The CES had already adopted a code of ethics a few years prior and had decided to contribute to the development and to adopt for itself the Joint Committee on Standards for Education Evaluation Programme Evaluation Standards (Yarbrough et al., 2012). The programme was launched in 2010 after substantial efforts were expanded to craft and organise the required systems, policies, administrative procedures, governance and management processes (Kuji-Shikatani et al., 2015).

Between July 2010 and February 2019, 428 individuals were granted the Credential Evaluator designation. The CES maintains an online list of these individuals. Through its lifespan, the programme has steadily gained recognition in Canada (Lahey et al., 2018). In 2015, the CES commissioned an independent evaluation of the programme (Fierro et al., 2016, p. 5) which concluded that the programme “is making strides in achieving several of the near-term intended outcomes. However, these achievements

sit against a backdrop that indicates continued progress may be at risk. Throughout the report we point to several areas where improvements can be made, and we hope that a thorough review of this document will help CES to improve several processes. Following are some specific recommendations about the most pressing issues that need attention to facilitate the success and sustainability of the current program”. Jean A. King (2015, p. 134) wrote “The Credentialed Evaluator designation has provided proof of concept for a viable evaluator credentialing system run by a voluntary organisation of professional evaluators (VOPE). Specific considerations in moving forward in settings beyond Canada include the following: (a) the exercise of caution when using evaluator competencies to structure a credentialing program, (b) the importance of a perceived need for or value of a credential, (c) skillful attention to milieu, (d) finding qualified and committed people to develop and manage the program, and (e) ensuring that all stakeholders, including those outside the profession, are involved”.

Lahey et al. (2018) have recently studied the demand and supply sides of the evaluation industry in Canada. They confirmed that the federal government remains the key player on the demand side—if only because its demand is more standardised or monolithic than that of other sectors that are dispersed organisationally and in the nature of their needs. They approximate the composition of the supply of evaluators in Canada from a 2016 survey of 405 CES members where “One-half of respondents considered themselves to be producers of evaluation results for their own organisation (50 percent) while a similar proportion were producers of evaluation results for other organisations (49 percent). One-quarter of survey participants said they were researchers on evaluation (23 percent) while one-fifth said they were users of evaluation results (21 percent)” (Probe Research, 2016, p. 10).

There is fluidity between employment in the private and public sectors, mainly with private sector evaluators getting hired in the federal public sector. The demand for internal evaluators in the federal government has translated into many young private sector evaluators moving to the federal public sector (Lahey et al., 2018). Given that the private sector is characterised by a small number of larger firms and a very large number of boutique suppliers (about one-half of companies comprising fewer than five employees), this dynamic has exercised pressures on the private sector.

A handful of academics provide practical evaluation services (as opposed to research on evaluation). Their work tends to serve needs of

NFPs (e.g. community-based organisations, social services organisations, international development organisations), provincial or municipal governments, and, to a lesser extent, federal buyers of evaluation (Lahey et al., 2018).

While the profession of evaluation is well structured in Canada, there is no organisation that acts as an ‘arbitration board’ to which someone could turn if they were dissatisfied with an evaluation study or an evaluator’s behaviour.

### *Compliance to Standards and Quality Obligations*

The CES has developed a code of ethics in the 1990s (CES, no date); it is currently revising it and a new approach to ethics should be promoted in 2020. New and renewing members of the CES must check a box on the membership form to indicate that they have ‘read and will adhere to the Canadian Evaluation Society’s Standards of Ethical Guidelines’. There are no other obligations made to CES members or mechanisms to oversee or enforce the code of ethics.

The CES has adopted the Evaluation Standards developed by the Joint Committee on Standards for Educational Evaluation (Yarbrough et al., 2012) as its own. In fact, the CES has been one of the standing members of the Joint Committee for over a decade. New and renewing members of the CES must check a box on the membership form to indicate that they have “read and will adhere to the CES Evaluation Standards (developed by the Joint Committee on Standards for Educational Evaluation)”. The standards act as suggestions rather than rules: they are not enforced and verification is not made that they are in fact used.

The Treasury Board of Canada Secretariat has adopted a standard for evaluation practice (Appendix 3 in TBS, 2016b). These standards concern how evaluations are planned and executed, who is involved in the evaluation process, and what evaluation reports should contain. These standards are binding for federal evaluators. Their respect can be the subject of a neutral assessment every five years; such an assessment is required according to the federal Policy (TBS, 2016a, paragraph 4.3.19).

The OAG periodically audits the evaluation function of the Government of Canada. There were such audits reported in 2009 (OAG, 2009) and 2013 (OAG, 2013). In 2009, the OAG expressed concerns about how evaluations were being conducted in the departments and about the role of the Secretariat. The main issues were the lack of availability of

performance information, the limited capacity of departments to meet the requirements for evaluation and the low level of use of evaluation findings to support programme improvement and expenditure management. In 2013, the OAG observed progress since 2009 in the capacity of departments to evaluate and on the processes in place to increase the use of evaluation findings and recommendations. Moreover, some departments had improved capacity in the area of performance measurement. The OAG concluded that there were still weaknesses in the contribution of programme evaluation to decision-making.

## CONCLUSION

Evaluation in Canada is driven by the federal government sector. This federal drive has evolved with highs and lows since 1977 since the introduction of the first Policy on Evaluation which was contemporaneous with the development of results-based management. While old in comparison with many other countries, the federal model of central oversight and ministerial responsibility (where government departments perform evaluations) has remained constant, though the policy has changed on three times; the 2016 Policy on Results subsumes evaluation matters and constitutes the fourth policy on evaluation. The federal policy is accompanied by standards of practice.

Evaluation exists in other spheres like provincial governments and the not-for-profit sector, in part as a response to the requirements of funders, that is, the federal government in the case of provinces and foundations in the case of not-for-profits.

Evaluation contributes to a number of public policy and programming decisions, mainly at the federal level but also in some provinces, but it has not found its way into the public debate. References to evaluation studies are very rare in the media and in civil society discussions.

The Canadian Evaluation Society is the oldest professional association of evaluators in existence globally. It enjoys a membership of approximately 1,800. It serves as an important focal point in providing a network for evaluators in Canada and important professional development for evaluators. The Society launched the Credentialed Evaluator professional designation in 2010; while improvements are always possible, it is considered a success at shaping the profession and contributing to a sense of belonging that was not as strongly expressed before its introduction.

Academic training in evaluation is available across Canada. Specialised graduate training is offered in the three largest provinces and available online.

Future challenges include maintaining relevance in the face of appetite for rapid performance information, broadening the environments where evaluation is routinely performed, contributing to debates on public policy, accumulating knowledge over the long term, and incorporating environmental and social sustainability concerns.

### LIST OF ABBREVIATIONS

CEE	Centre of Excellence for Evaluation
CES	Canadian Evaluation Society
MRAP	Management Response and Action Plan
NFP	Not-for-Profit
OAG	Auditor General of Canada
TBS	Treasury Board of Canada Secretariat

### REFERENCES

- OAG. (2009). *Fall report of the auditor general of Canada: chapter 1—Evaluating the effectiveness of programs*. [http://www.oag-bvg.gc.ca/internet/English/parl\\_oag\\_200911\\_01\\_e\\_33202.html](http://www.oag-bvg.gc.ca/internet/English/parl_oag_200911_01_e_33202.html). Accessed on 9 November 2018.
- OAG. (2013). *Spring report of the Auditor General of Canada: chapter 1—Status report on evaluating the effectiveness of programs*. [http://www.oag-bvg.gc.ca/internet/English/parl\\_oag\\_201304\\_01\\_e\\_38186.html](http://www.oag-bvg.gc.ca/internet/English/parl_oag_201304_01_e_38186.html). Accessed on 7 February 2019.
- Buchanan, H. (2015). A made-in-Canada credential: Developing an evaluation professional designation. *Canadian Journal of Program Evaluation*, 29(3), 33–53.
- Canadian Encyclopedia. (2013). *British North America Act*. <https://www.thecanadianencyclopedia.ca/en/article/british-north-america-act>. Accessed on 8 February 2019.
- Canadian Encyclopedia. (2015). *Land*. <https://www.thecanadianencyclopedia.ca/en/article/land>. Accessed on 7 February 2019.
- CES. (1980). *Proceedings from the first annual conference of the Canadian evaluation society*. [https://evaluationcanada.ca/txt/c1980\\_proceedings.pdf](https://evaluationcanada.ca/txt/c1980_proceedings.pdf). Accessed on 31 January 2019.

- CES. (1981). *Letters patent* [https://evaluationcanada.ca/txt/CES\\_Letters\\_Patent\\_1981.pdf](https://evaluationcanada.ca/txt/CES_Letters_Patent_1981.pdf). Accessed on 31 January 2019.
- CES. (no date). *Ethics*. <https://evaluationcanada.ca/ethics>. Accessed on 7 February 2019.
- CEE. (2015). *Evaluation of the 2009 policy on evaluation*. <https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/centre-excellence-evaluation/evaluation-2009-policy-evaluation.html>. Accessed on 8 February 2019.
- Chouinard, J. A., & Milley, P. (2015). From new public management to new political governance: Implications for evaluation. *Canadian Journal of Program Evaluation*, 30(1), 1–22.
- Docherty, D. C. (2014). Parliament: making the case for relevance. In J. Bickerton & A.-G. Gagnon (Eds.), *Canadian politics, sixth edition*. University of Toronto Press.
- Doern, G. B. (2007). Innovation, science, environment 06/07. *Canadian Policies and Performance, 2006–2007*. McGill-Queen's University Press.
- Dumaine, F. (2012). When one must go: The Canadian experience with strategic review and judging program value. In G. Julnes (Ed.), *Promoting valuation in the public interest: Informing policies for judging value in evaluation—New Directions for Evaluation*, No. 133. (pp. 65–75). American Evaluation Association.
- Fetterman, D. M. (2005). In response to Drs. Patton and Scriven. *American Journal of Evaluation*, 26(3), 418–420.
- Fierro, L. A., Galport, N., Hunt, A., Codd, H., & Donaldson, S. I. (2016). *Canadian evaluation society credentialed evaluator designation program evaluation report*. Claremont Graduate University.
- Gauthier, B., Barrington, G. V., Bozzo, S. L., Chaytor, K., Dignard, A., Lahey, R., Malatest, R., McDavid, J. C., Mason, G., Mayne, J., & Porteous, N. L. (2004). The lay of the land: Evaluation practice in Canada today. *Canadian Journal of Program Evaluation*, 19(1), 143–178.
- Gauthier, B., Barrington, G. V., Bozzo, S. L., Chaytor, K., Dignard, A., Lahey, R., Malatest, R., McDavid, J. C., Mason, G., Mayne, J., & Porteous, N. L. (2009). The lay of the land: Evaluation practice in Canada in 2009. *Canadian Journal of Program Evaluation*, 24(1), 1–49.
- Greene, I. (2002). Lessons learned from two decades of program evaluation in Canada. In D. Braunig & P. Eichorn (Eds.), *Evaluation and accounting standards in public management* (pp. 44–53). Nomos Verlagsgesellschaft.
- Halpern, G., Gauthier, B., & McDavid, J. C. (2015). Professional standards for evaluators: The development of an action plan for the Canadian evaluation society. *Canadian Journal of Program Evaluation*, 29(3), 21–32.
- Hunter, T., & McDavid, J. C. (2018, October). *Inventory of Canadian graduate evaluation education*. Consortium of Universities for Evaluation Education.



- International Monetary Fund. (2018). *World economic outlook database*.
- Jacob, S. (2006). Trente ans d'évaluation de programme au Canada : l'institutionnalisation interne en quête de qualité. *Revue Française D'administration Publique*, 119, 515–532.
- Jacob, S., & Speer, S., & Furubo, J.-E. (2015). The institutionalization of evaluation matters: Updating the international atlas of evaluation 10 years later. *Evaluation. The International Journal of Theory, Research and Practice*, 21(1), 6–31.
- King, J. A. (2015). From the outside, looking in with a smile: A summary and discussion of CES's credentialed evaluator designation. *Canadian Journal of Program Evaluation*, 29(3), 134–153.
- Kuji-Shikatani, K., Thompson, M., & Matthew, M. (2015). Launching the credentialed evaluator (CE) designation. *Canadian Journal of Program Evaluation*, 29(3), 70–85.
- Lahey, R. (2010). *The Canadian M&E system: lessons learned from 30 years of development*. Evaluation Capacity Development (Working Paper Series No. 23). World Bank: November 2010.
- Lahey, R. (2011). *The Canadian monitoring and evaluation system*. PREM Notes No. 11, Special Series on the Nuts and Bolts of Government M&E Systems. World Bank: June 2011.
- Lahey, R. (2015). *Strategic-level review and the role of evaluation. An insider's view. Keynote presentation to the annual conference of the Canadian Evaluation Society*. Montreal, Québec.
- Lahey, R. (2017, January). Connecting the dots between M, E, RBM and deliverology. *Canadian Government Executive*, 23(1), 6–8.
- Lahey, R., & Nielsen, S. B. (2013). Rethinking the relationship among monitoring, evaluation and results-based management: Observations from Canada. *New Directions for Evaluation*, 137(Spring), 45–56.
- Lahey, R., Elliott, C., & Heath, S. (2018). The evolving market for systematic evaluation in Canada. *New Directions for Evaluation*, 160, 45–62.
- Lahey, R. et al. (forthcoming). *Evaluation in the provinces and territories. A cross-Canada review*.
- Lasby, D. (2018). *The state of evaluation measurement and evaluation practices in Ontario's nonprofit sector*. Ontario Nonprofit Network.
- Lasby, D. (2019). *The state of evaluation: measurement and evaluation practices in Canada's charitable sector*. Imagine Canada.
- Love, A. (2015). Building the foundation for the CES professional designation program. *Canadian Journal of Program Evaluation*, 29(3), 1–20.
- Maicher, B., & Frank, C. (2015). The development and initial validation of competencies and descriptors for canadian evaluation practice. *Canadian Journal of Program Evaluation*, 29(3), 54–69.

- Probe Research. (2016). *General national survey of CES membership prepared for the Canadian Evaluation Society*. [https://evaluationcanada.ca/txt/201701\\_ces\\_survey\\_report\\_en.pdf](https://evaluationcanada.ca/txt/201701_ces_survey_report_en.pdf). Accessed on 2 February 2019.
- Savoie, D. (2014). Power at the apex: Executive dominance. In J. Bickerton, & A.-G. Gagnon (Eds.), *Canadian politics, sixth edition*. University of Toronto Press.
- Statistics Canada. (2018a). *Canada's population estimates, third quarter 2018*. <https://www150.statcan.gc.ca/n1/daily-quotidien/181220/dq181220c-eng.htm?HPA=1&indid=4098-1&indgeo=0>. Accessed on 7 February 2019.
- Statistics Canada. (2018b). *Labour force characteristics by industry, annual*. <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410002301>. Accessed on 7 February 2019.
- Thomas, P. G. (2014). Two cheers for bureaucracy: Canada's public service. In J. Bickerton, & A.-G. Gagnon (Eds.), *Canadian politics, sixth edition*. University of Toronto Press.
- TBS. (no date). *Database on program evaluation tracking*. Government of Canada, unpublished.
- TBS. (2000). *Results for Canadians, a management framework for the government of Canada*. [https://www.tbs-sct.gc.ca/report/res\\_can/rc-eng.pdf](https://www.tbs-sct.gc.ca/report/res_can/rc-eng.pdf). Accessed on 14 February 2019.
- TBS. (2008). *Policy on transfer payments*. Government of Canada.
- TBS. (2012). *Policy on transfer payments*. Government of Canada.
- TBS. (2016a). *Policy on results*. Government of Canada. <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=31300>. Accessed on 31 January 2019.
- TBS. (2016b). *Directive on results*. Government of Canada. <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=31306>. Accessed on 7 February 2019.
- Yarbrough, D. B., & Shulha, L. M., & Hopson, R. K., & Caruthers F. A. (2012). *The program evaluation standards* (3rd ed.). Sage.



## Evaluation in Chile

*Claudia Olavarría Manríquez and Andrea Peroni Fiscarelli*

### GENERAL OVERVIEW: INSTITUTIONALITY OF EVALUATION<sup>1</sup>

Chile is a country with a presidential regime, a bicameral system in the legislative branch and an independent judiciary. The country relies on a modern public administration system that is governed by the concepts of results-based management.

<sup>1</sup> We extend our thanks to those who collaborated in the elaboration of this article by giving their vision and time: to the Deputy for the 6th District, Valparaíso Region, Diego Ibañez; to the Head of the Management Control Division of the Budget Office Paula Darville; to the Head of the Social Policy Division of the Ministry of Social Development Amanda Dawes; to the Coordinator of the Law Evaluation Department of the Chamber of Deputies Maryan Henríquez, to the President of the Civil Society Council of the Ministry of Social Development Nicole Romo, to the (former) Coordinator of the Chilean Evaluation Network, EvalChile Dominique Keim, to the academic from the University of Chile and former Undersecretary of Evaluation of the Ministry of Social Development Heidi Berner and to the Director of the Public Policy Centre of the Pontificia Universidad Católica de Chile Ignacio Irrazábal.

---

C. Olavarría Manríquez (✉)

Independent Evaluation Consultant, Santiago, Chile

A. Peroni Fiscarelli

Universidad de Chile, Santiago, Chile

e-mail: [aperoni@uchile.cl](mailto:aperoni@uchile.cl)

In search of adding transparency and better results regarding public spending, the Chilean State began a process of adopting Monitoring and Evaluation (M&E) systems since 1990, after the recovery of democracy. The progressive implementation of these M&E systems was accompanied by a series of initiatives aiming to modernise public management. This generated a favourable environment allowing its emergence as strategic management exercises, as well as the introduction to the use of information and communication technologies, the installation of offices for information delivery and receipt of complaints, the introduction of new mechanisms of performance incentives, the creation of a system of senior public management and the existence of rules on transparency and integrity (Irrazaval & De los Rios, 2015). According to the assessment carried out, in 2007, by the Latin American Centre for Development Administration (CLAD) and the Inter-American Development Bank, Chile has managed to install perhaps the most advanced experience in terms of monitoring systems and follow-up in Latin America. It has incorporated results-based management tools in a systematic and methodical way, having tested and improved various instruments<sup>2</sup> during this experience (Olavarría & Peroni, 2012).

There are currently two M&E systems in the country, which jointly cover government action (Rios, 2007). These are located in the Ministry of Social Development and Family (MDSF), in the Undersecretary of Evaluation, and in the Ministry of Finance, in the Budget Office (DIPRES). Since the 1960s, investment evaluations have been carried out in an institutionalised manner in the National Planning Office (ODEPLAN), on the other hand, the “ex post evaluations [as they are currently developed] began in 1997 and from 2003 the DIPRES is granted the power to carry out evaluations of social, productive development and institutional development programmes” (DIPRES, 2019, p. 4).

Having said this, the “installation of these systems, [...] evidences the advances regarding the incorporation of monitoring and evaluation instruments and practices in the daily endeavour and organisational culture of the Chilean State, with a close connection to the concepts of

<sup>2</sup> The Management Control System working with Strategic Definitions, the Management Improvement Programme, Performance Indicators, Integral Management Balance and Competitive Funds, among others, illustrate this.

accountability and the generation of quality public expenditure (resource allocation and investment selection)” (Olavarría & Peroni, 2012, p. 5).

## INSTITUTIONAL STRUCTURES AND PROCESSES (POLITICAL SYSTEM)

### *Evaluation Regulations*

The first regulation referring to the role of the State as an overseer of both the good use of public resources and their evaluation, is found in the Decree Organic Law of the State Financial Administration, Number 1,263 of the Board Government of the Republic of Chile, Santiago, November 21, 1975, having agreed as follows: “Article 53: The Office of the Comptroller General of the Republic may request to the public services subject to its audit the necessary reports allowing the verification of the income and expenses corresponding to their management”. After that, the regulation is specified regarding the ministries in charge of evaluation, according to the laws that created them, their attributions and the roles they fulfill in this regard.

#### **A. Ministry of Social Development and Family (MDSF), Under-secretary of Evaluation**

The MDSF was created in 2011, following the orientation of ODEPLAN<sup>3</sup> (1967) and after the Ministry of Planning and Cooperation (as from 2018 the MDSF included the term Family and became the MDSF). The MDSF and the Under secretariat of Social Evaluation were created following Law No. 20,530. The specific task of the Under secretariat of Social Evaluation is the design, coordination and evaluation of the Government’s social policies, in order to contribute to the improvement of social spending focus through the permanent evaluation of programmes implemented by the State. It is also responsible for the analysis and technical–economic evaluation of public investment initiatives.

<sup>3</sup> The mission of ODEPLAN was to promote and organise the planning system at national, sectoral and regional levels. See: <https://www.casamuseoeduardofrei.cl/objeto-del-mes-nacimiento-de-la-oficina-de-planificacion-nacional/>. Accessed on 03 November 2020.

In the MDSE, two evaluation subsystems are carried out, (i) the Integrated Bank of Social Programmes (BIPS) and (ii) the Integrated Bank of Projects (BIP) of the National Investment System (SNI). The BIPS is a record that contains information corresponding to those social programmes, whether ongoing or not, that have been or are being submitted to any of the evaluations referred to in letter (c) and letter (d) of article 3 (Article 2, Law No. 20,530). The BIP is a registry that contains those investment initiatives that have been evaluated, whether ongoing or not, and require State financing. The evaluation tasks established by law are extended by carrying out a consultation process with the Civil Society Council and by establishing recommendation actions oriented towards the State (Article 3).

#### **B. Ministry of Finance, Budget Office (DIPRES)**

Within the framework established in Decree Law, 1,263 (1975), in article 52, mentioned above, and its Regulation N° 1177 (2003), the DIPRES is granted the legal attribution to carry out annual evaluations of social, productive and institutional development programmes included in public service budgets. In this context, the Regulation of the Law (Darville et al., 2017) establishes, among others: (i) basic principles (independence, transparency); (ii) instances and institutions (Interministerial Committee, Budget Office, Evaluated Institutions), with specific tasks; and (iii) types of evaluations.

The Management Evaluation and Control System is located in the Public Management Control Division of the DIPRES. The aforementioned system addresses the areas of budgetary management, improvement of management of institutions and governance. It aims at maintaining quality public spending. This system allows to rely on information on the performance of public institutions to support decision-making during the different stages of the budget process, incorporating the concept of ‘budgeting for results’.

In other Ministries or Public Services of the State of Chile, the exercise of evaluation depends on its priorities, therefore such task is not established, by means of a specific regulation.

### *Evaluation Practice*

The practice of state evaluation differs according to each ministry and, in turn, it is distinguished by the course of action that each of them develops.

The BIPS (MDSF) makes the official list of social programmes and a description for each of them available to the public. In the case of social programmes that are being implemented by the State, a follow-up report is additionally included. In the case of new and significantly reformulated programmes, a recommendation report on the results of the ex ante evaluation is also incorporated. The report of recommendations contemplates the evaluation, among others, of the consistency, coherence and compliance of such social programmes. This analysis will be a factor to be considered regarding the coordination of the public offer and it will also be contemplated concerning resource allocation, as the draft Budget Law is formulated.

Through the BIPS, the Ministry is committed to informing citizens about social programmes and, in this way, increasing transparency about initiatives and the use of public resources related to them. To this end, information concerning the design and subsequent monitoring of each social programme is published.

To date, the programmes evaluated are the following in Table 6.1.

**Table 6.1** Social programmes evaluated ex ante, MDSF Chile (2012–2018) (Ministerio de Desarrollo Social y Familia 2019)

<i>Submission year</i>	<i>Programmes submitted for evaluation</i>	<i>Ministries</i>
2012	44	9
2013	55	10
2014	124	12
2015	87	11
2016	53	9
2017	56	12
2018	113	11
	532	

In relation to the Monitoring of Social Programmes, the MDSF carries out an annual monitoring concerning the management and implementation of ongoing social programmes, which are being implemented by public services related to or dependent on it and other ministries. In

2017, the MDSF monitored 333 Social Programmes in 13 Ministries, representing \$ 5,323,533 million (BIPS, 2018).

The SNI of the MDSF<sup>4</sup> regulates and governs public investment in Chile. It also brings together the methodologies, standards and procedures that guide the formulation, execution and evaluation of those Investment Initiatives applying to public funds. The SNI is composed of four subsystems, which define the stages of the investment process, namely: Ex-Ante Evaluation and Ex-Post Evaluation Subsystems, managed by the MDSF; and Budget Formulation and Budget Execution Subsystems, managed by the DIPRES. The BIP is an information system designed to support public investment management by registering the basic projects, programmes and studies that request financing annually. Methodologies applied refer to the economic evaluation of projects.

In the Ministry of Finance, within the DIPRES, the evaluation process clearly establishes tasks and roles regarding institutional actors involved: Interministerial Committee, Budget Office, Evaluated Institutions (Darville et al., 2017, p. 11).

- a. Interministerial Committee. It is formed by a representative of the Ministry—General

Secretariat of the Presidency, the MDSF and the Ministry of Finance, through the Budget Office, who chairs it.<sup>5</sup>

- b. Budget Office. It is responsible for the proper execution of evaluations.<sup>6</sup>

<sup>4</sup> See: Integrated Project Bank: <https://bip.ministeriodesarrollosocial.gob.cl>, accessed on 10 August 2019.

<sup>5</sup> This Committee aims to ensure that the carrying-out of evaluations is consistent with government policies; that conclusions arising from this process are known by the institutions that comprise it and; that the necessary technical support and coordination are available for its proper development, especially in the process of program selection and selection of consultants.

<sup>6</sup> It is especially up to it to carry out the process of selection of the programmes to be evaluated; definition of methodological and operational designs; provision of resources for its operation; management of the operation; analysis and approval of the progress and final reports indicating the observations that it deems appropriate; reception and submission of evaluation reports to the National Congress and public institutions; integration of evaluation results into the budget cycle and; establishment of institutional commitments to improve the programmes evaluated.



c. Evaluated Institutions. They participate in the process through the following activities:

(i) To provide information to be delivered to the evaluation team, at the beginning of the process, and any other information available and necessary during the evaluation process. (ii) To participate in all meetings requested by the evaluators and the Ministry of Finance, in order to analyse specific aspects during the evaluation process. (iii) To review and issue comments on the intermediate and final evaluation reports, to be delivered to the evaluators through the Ministry of Finance. (iv) To define the actions required in order to improve programme results by establishing institutional commitments with the Budget Office based upon the recommendations made in the evaluation. (v) To inform the Budget Office of the fulfillment of institutional commitments, semiannually and during the process of annual budget formulation, in accordance with the instructions issued by the Ministry of Finance.

In this way, in order to contribute to decision-making, the selection of programmes to be evaluated arises from proposals made by the different actors of the Interministerial Committee.

Additionally, while the Budget Law is undergoing legislative proceedings, Congress can complement or suggest new evaluations according to the priorities of the Legislative Branch or according to requirements that arise during the budget discussion process. In order to ensure the independence of the institution evaluated, most evaluations are carried out by entities outside the public sector, whether they are consultants or universities of recognised prestige, selected through public bidding processes, or by panels of experts, selected via public contest. Regarding transparency, all final evaluation reports are sent to Congress and published on the DIPRES website. To ensure that evaluation results have effects and are translated into concrete improvement actions, either regarding programme management or the allocation and/or use of their resources, commitments are defined by the Services in charge of the programmes evaluated. These are prepared based upon the findings and recommendations of the evaluation and are agreed jointly between the Service and the DIPRES. (DIPRES, 2019, p. 4)

In this way, the actors involved and the operation process are clearly established, as well as the extent to which evaluation results contribute

to budgetary decisions, in Parliament. The evaluations developed by the DIPRES are the following:

- Evaluation of Government Programmes (EPG).<sup>7</sup>
- Evaluation of Institutional Expenditure (EGI).<sup>8</sup>
- Impact Evaluation (EI).<sup>9</sup>
- Focalised Scope Evaluation (EFA)<sup>10</sup> (DIPRES, 2019, p. 4).

The most used evaluation model is the EPG, according to which, as it is also the case in other evaluation types, “programmes evaluated are classified following a performance category, which is defined considering aspects regarding the results (intermediate and final) of the programme, its design, implementation and / or efficiency in the use of resources”<sup>11</sup> (DIPRES, 2019, p. 6). Performance classification is based upon findings concerning the four areas (design, implementation, efficiency and results) or some of them, as the case may be depending on the evaluation line. The DIPRES prepares an evaluative judgment<sup>12</sup> concerning each of the

<sup>7</sup> This line evaluates the consistency of the objectives and design of the programme, aspects of its organisation and management and results at the product level (coverage, targeting, among others).

<sup>8</sup> Its goal is to evaluate design and institutional management, as well as the results and use of resources regarding the supply of strategic products of the institution.

<sup>9</sup> The focus is on evaluating intermediate and final results. Experimental or quasi-experimental methodologies are used, and information is collected through surveys, focus groups and interviews, among others.

<sup>10</sup> Through this line, evaluations focus on specific areas of programme performance: (i) strategy design, (ii) costs and (iii) implementation. Due to the fact that evaluations are narrower and shorter, results are more appropriately delivered.

<sup>11</sup> Each of these areas is defined as follows:

- Design: verifies the validity of the problem identified and the intervention strategy of each programme.

- Implementation: refers to the analysis of the focus of the programme, as well as the identification of gaps between the theoretical design and its execution.

- Efficiency: refers to an evaluative analysis of the cost of producing the goods or services.

- Results: verifies results at purpose level, in relation to changes in the situation of beneficiaries as an effect of the delivery of the goods or services of the programme (DIPRES 2019, p. 6).

<sup>12</sup> On each relevant area, the DIPRES prepares an evaluative judgment (Sufficient, Insufficient, Not Concluding). Based on these judgments, programmes are classified into

four areas, although in general, it considers intermediate results, since the EPG, due to its characteristics, does not always evaluate final results. Once results are published, the programmes evaluated are informed. They formally respond to the conclusions and recommendations of the evaluation. This information is presented to Parliament and it is considered for the next budget allocation process, installing a system for monitoring the commitments assumed by the institution evaluated.

To date, the programmes evaluated are the following shown in Table 6.2.

At budget level, coverage reaches its highest point in the 2005–2010 period, with 18.4% of the evaluable budget (see Table 6.3).

The ex post evaluation system carried out by the DIPRES has evaluated a total of 533 programmes and 50 institutions in the period 1997–2017. Most evaluations correspond to the EPG line, which accounts for almost 70% of total evaluations. Second, it is followed by impact evaluations (EI) that concentrate 20%. The remaining 10% is distributed on the EGI, Evaluation of New Programmes and EFA lines. 23% of the total existing institutions in 2017 have been evaluated<sup>13</sup> (Darville et al., 2017, p. 36). In terms of programmes, the number of programmes evaluated in the 2012–2017 period represents about 20% of the total evaluable programmes currently available (2017). On average, 4.9% of the budget has been evaluated annually in the period 1999–2017; a percentage that goes from 0.8% in 2012 to 7% in 2017. In absolute terms, evaluation efforts have focused on the Ministries of Health, Labor and Social Welfare, Education and Housing and Urban Planning, which correspond

four performance categories that also consider seniority as a proxy for the level of maturity of the programme: Good Performance, Medium Performance, Low Performance and Bad Performance (DIPRES 2019, p. 6).

<sup>13</sup> Regarding the coverage by ministry (1999–2017), the only ministry having carried out annual evaluations during the entire period analysed is the Ministry of Education. On average, the annual coverage reaches 2.7%, with a maximum of 9.9% in 2011. The average coverage of the Ministry of Health is 8.4%, with a maximum of 88.9% in 2011 for the evaluation of the National Health Fund. In the Ministry of Labor and Social Security, the average coverage is 6.1%, with a maximum of 88.7% in 2007, concerning the evaluation of the National Welfare Institution. Other ministries that have had evaluations in almost every year of the period analysed are the Ministry of Interior and Public Security and the Ministry of Social and Family Development. In the first case, in 17 of the 19 years analysed, evaluations are observed and the annual average coverage is 3.8%. In the second case, evaluations are observed in 18 years and the average annual coverage is 13.8%. (Darville et al., 2017, pp. 30f.).

**Table 6.2** Number of institutions evaluated and coverage. Period 2002–2017 DIPRES (Darville et al., 2017, p. 24)

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2013	2015	2016	2017
Institutions evaluated	6	2	4	1	2	4	7	5	5	3	6	2		
Institutions to be evaluated													1	2
Total	6	2	4	1	2	4	7	5	5	3	6	2	1	2
Evaluable institutions	171	175	177	189	190	190	194	196	196	199	202	203	204	204

**Table 6.3** Budget evaluated. Coverage (1999–2017) (Darville et al., 2017, p. 28)

<i>Seven year period</i>	<i>Evaluated</i>	<i>Evaluable</i>	<i>Coverage (%)</i>
P1 (1999–2004)	\$505.230	\$16.000.135	3,2
P2 (2005–2010)	\$5.034.716	\$27.394.764	18,4
P3 (2011–2017)	\$5.562.878	\$39.781.879	14,0

precisely to the ministries with the highest evaluable budgets. In relative terms, evaluation efforts vary. In the first period, 1999–2004, the Ministries of Agriculture and Social Development stand out. During the 2005–2010 period, the coverage of the Ministries of Justice, Housing and Urban Planning and Labor and Social Welfare stands out, exceeding 80%. For the last period (2011–2017), rates above 80% in the Ministries of Health and Social Development stand out (Darville et al., 2017, p. 36).

In the other Ministries or Public Services, evaluations carried out in general terms are outsourced through public bidding and carried out by external teams. This is due, on the one hand, to the limited capacity of the teams working in the Departments of Studies, Planning and Evaluation, to assume new tasks, and on the other hand, it responds to the need for it to refer to the principle of independence (of designers or implementers).

Regarding the Ministry of the Environment, the development of the Environmental Impact Assessment System (SEIA) is worth noting. The SEIA entered into force on April 3, 1997. “This instrument allows the introduction of the environmental dimension in the design and execution of projects and activities carried out in the country; through it, it is evaluated and certified that the initiatives, both of the public and private sectors, are in a position to comply with the environmental requirements that apply to them” (Servicio de Evaluación Ambiental, 2019). More than two decades after its application, more than 24,700 projects or activities have been approved in the SEIA, which has allowed the country to be able to evaluate the impact that public and private investments can generate, or to mitigate significant adverse impacts (Servicio de Evaluación Ambiental, 2019).

The Legislative Branch carries out evaluations within the framework of a programme called Law Evaluation. This programme, arisen from concerns of parliamentarians, performs periodic evaluations concerning a sample of legislative initiatives, in order to know how the operation of

the law has been applied. The objectives of the programme are: to determine the degree of compliance with the objectives expected as the law was adopted, to identify externalities, impact and undesired consequences, to learn about citizens' perception of the law and its implementation and to propose corrective measures regarding the law and its implementation (OECD, 2012).

Law evaluations are ex post evaluations, carried out by a National Congress internal evaluation team, which applies a three-step standard methodology, with emphasis on citizen participation, namely: technical analysis of the law, survey and citizen perception analysis and preparation of a final report (OECD, 2012). This team has carried out two to three law evaluations, per legislative period since its creation, in 2010.

### *Use of Evaluations*

As noted, M&E systems institutionalised in the State of Chile are implemented with the main purpose of contributing to the quality of public spending and along these lines, the results of these systems are effectively integrated into the budget cycle, improving efficiency in the allocation and use of public resources, as well as quality of expenditure and institutional management (DIPRES, 2019).

The use of evaluation by Parliament is then evidenced by the processing and approval of the Annual Budget Law, for which the results of programme evaluation, both ex ante and ex post, constitute a fundamental input. There is no consensus among the key actors interviewed on the depth of the analysis and use of evaluations carried out made by Parliament. According to Parliament, even though a debate based upon technical and evidence-based nature should be strengthened in order to delimit the Budget Law, many of the definitions made in this space respond to political criteria. On the other hand, people interviewed having experience in the operation of M&E systems show how a more extensive use of the evaluation results is observed. Parliamentarians use not only the results of ex post evaluations, but also ex ante evaluation and monitoring reports. Despite these differences, the need to strengthen the evidence for decision-making in all parliamentary teams, not only in ad hoc thematic commissions, is observed as transversal.

This use made by the Parliament is complemented by the use of results deriving from Law evaluation initiatives, which, although smaller, have the task of strengthening initiatives arising from the Congress.

Regarding the use of evaluation in the various sectors of the State, the following should be noted:

- Each Ministry under secretariat or Service orders evaluations according to its needs to generate evidence, these evaluations are tendered to external teams. The results of these evaluations are used, in general terms, by the various sectors for the improvement of their programmes and/or to comply with management requirements.
- Although the main purpose of the aforementioned M&E systems is to provide evidence for budgetary decision-making, they also contemplate feedback and improvement of the evaluated initiatives. An example of this are improvement commitments agreed among the DIPRES and the teams of the institutions evaluated. These commitments derive from the recommendations of evaluations carried out, they refer to various areas for improvement of the design and/or implementation of the programmes evaluated and must be implemented by internal teams. The DIPRES tracks its implementation. The perception of key actors linked to the management of the M&E system shows a heterogeneous performance between sectors, which can be explained because of various reasons, such as political will, technical competence of the teams, the quality of the data collected, the complexity of the management structures, among others.

There is no systematised and available information on the use of evaluation at sector level, however, the perception of the key stakeholders consulted gives light on these differences. Respondents recognise a heterogeneous performance in terms of evaluation processes and use of results concerning various State sectors.

It is to be observed that programmes in the social sector show good performance concerning the adoption of evaluation practices and the consequent use of their results. Social programmes under the Ministry of Social Services incorporate M&E practices in their intervention strategies and use results to improve targeting and redesign strategies, among others. At this point it is important to highlight that the MDSF has implemented monitoring strategies regarding social programmes evaluated in their *ex ante* stage. These practices are fully adopted by the internal teams.

Housing and public infrastructure programmes are mentioned as programmes showing good performance regarding incorporation of evaluation practices and use of the results thereof, particularly because of the capacities installed concerning the use of social evaluation tools for investment projects.

The economic sector, as well as the education sector are highlighted as variable performance sectors, which is explained by a multiplicity of factors such as expertise concerning teams, complexity of programme management strategies and the heterogeneous level of internalisation of monitoring practices, among others.

The key actors interviewed refer to the health sector as an area of diminished performance. They recognise to what extent the evaluation of programmes implemented in this sector represents a challenge for the evaluation teams due to various factors, among them the wide coverage and the management model of programmes, the quality of records containing monitoring data, the expertise of the teams and cultural aspects, among others.

Although there is not enough published information about the groups that use evaluation in Chile, it is clear from the actors' discourse that the main use of evaluation results by professional groups is linked to:

- Decision-making teams on budgetary allocation (parliamentarians and parliamentary advisors), which is the main objective of the systems reviewed.
- Teams implementing policies, programmes and projects and using evaluation as a tool to improve programme implementation.
- Teams designing policies, programmes and projects that use evaluation as 'lighting' of the aspects to be corrected and as evidence for the design of new initiatives.
- Academic teams, consultants and researchers using evaluation as evidence for the development of studies in various fields.

It is not possible to identify the existence of strategies used to guarantee the usage of evaluation.



## DISSEMINATION IN SOCIETY AND ACCEPTANCE (SOCIAL SYSTEM)

### *Institutionalised Use of Evaluations by Civil Society*

As reviewed during the development of this article, evaluation in Chile is mostly linked to public management and it is used for decision-making concerning the allocation of resources and the management of programmes and projects. Civil society has timidly begun to envision evaluation as a public good for advocacy or accountability.

In Chile, decision-making practices concerning referendums or community-based policies are not common and even less so is the use of evaluation as the basis for this type of decision-making. The limited existence of community-based decision-making practices in the public sphere is mainly due to the fact that programmatic proposals and social intervention projects arise from public bodies, based on methodologies in which community participation is not encouraged as a priority area. Although there are experiences of decision-making based on community-based referendums at municipal level, it is an exceptional practice that moves away from mainstream decision-making within the political system installed in Chile since the return to democracy.

Although civil society has not institutionalised or extended the use of evaluation results, there are some initiatives in this direction:

- There is a limited group of civil society organisations, larger and with a greater degree of professionalisation, implementing social programmes, with various accents, who conduct evaluations and disseminate their results.
- The evaluation of these results has diverse objectives, but accountability, incidence and dissemination of their work are mostly acknowledged. Some examples of these organisations are *Hogar de Cristo*, *Fundación para la Superación de la Pobreza* and *Techo*.
- In Chile, many civil society organisations are also public policy—mainly social policy—implementers, since they implement social programmes. By mandate of the State, these organisations implement monitoring and evaluation practices regarding their work, which are used for accountability and institutional learning. Organisations implementing programmes for the National Service for

Minors, some of which have advanced in the adoption and promotion of M&E systems of their endeavour with a rights approach such as the Option Corporation or the Chilean Pro United Nations Association, illustrate this.

- There are experiences of civil society organisations and associations of organisations that promote and facilitate the development of individual and also institutional evaluation capacities. In these organisations, strategies for advocacy and evaluation promotion are implemented in conjunction with other actors in the (eco) evaluation system. Some of these organisations emerge from universities such as the Public Policy Centre of the Pontificia Universidad Católica de Chile, and the Chilean Evaluation Network (EvalChile).
- EvalChile—which will be reviewed in the next chapter—is a Voluntary Organisation for Professional Evaluators (VOPE) that has implemented initiatives aimed at articulating civil society actors in order to promote and strengthen evaluation in Chile since 2009.

An interview held with a representative of organised civil society, carried out within the framework of this study, shows the need for both individual and institutional capacity building concerning M&E, through training strategies and the provision of spaces where lessons learned can be shared and articulation with various actors, both public and private, can be promoted. The above is especially important in the case of small and medium organisations.

A frequent use of evaluation results by individual citizens and/or private companies is not observed. Notwithstanding the foregoing, there is an incipient demand for developing social impact evaluations regarding Corporate Social Responsibility strategies—or similar—of private companies. Although it is not an extended practice, an incipient demand has been observed in recent years, leading to the development of consulting companies whose work focuses on this group.

The incorporation of civil society organisations in evaluations is not a common practice in institutionalised M&E systems in Chile. Nonetheless, since 2014 the MDSF has established a Civil Society Council, formed by representatives of the Ministry's services, civil society organisations, experts and academics. This council aims to incorporate the visions of its members, into the processes of discussion and decision-making on the design, execution and evaluation of public policies implemented by the Ministry (MDSF, 2019). As noted, both in the review of institutional web

pages as well as in the interviews carried out, there is an institutionalised space where the State encourages citizen participation in the evaluation of public policies, however, the level of involvement in decision-making in this regard is still incipient and it varies in relation to the emphasis given by the various administrations.

The limited incorporation of civil society into institutionalised M&E systems in the Chilean state is due to the fact that the design of these systems and the methodologies they implement do not contemplate an active or committed participation of civil society as part of their operation.

Beyond the aforementioned idea, the involvement of civil society in the evaluation process of various programmes could be acknowledged mainly in two moments: as beneficiaries in those evaluations that incorporate the collection of data on user perception—as satisfaction—and as recipients of the dissemination of evaluation results. On the last point, it should be clarified that this dissemination of results is carried out almost exclusively through the publication of reports on web pages without a greater focus on the dissemination strategies of results, which limits their scope and potential use.

Despite this limited incorporation of civil society in institutionalised State M&E systems, some examples of incorporation of these actors in methodological strategies and dissemination of results of evaluations implemented by other State institutions and by civil society organisations themselves, which implement evaluations because of its greater scope and professionalisation of their work, are observed.

Some examples of this trend are:

- As for the Legislative Branch, a fundamental part of the methodology of the Law Evaluation programme relies on the consultation of relevant civil society actors concerning their perception on law implementation (OECD, 2012).
- As for civil society organisations, there is an incipient trend towards the incorporation of participatory evaluation methodologies. Participatory evaluation initiatives developed by the *Fundación de la Superación de la Pobreza* and by *Techo* in the last period illustrate this (Sánchez & Miranda, 2019; Techo, 2019).

Despite this trend, an obstacle for its adoption is the overrating of quantitative and/or econometric methods and approaches for evaluation, at the expense of qualitative, participatory methods and approaches, among others. This has resulted in overrating, for example, of impact evaluations, which use experimental or quasi-experimental models, as a single way to account for the results obtained by programmes and projects, highlighting the potential of other methodological proposals.

As Olavarría and Peroni point out, there is a need to “Promote the increasing incorporation of participatory evaluation methodologies in monitoring and evaluation systems, as a mechanism to systematically record the perception of actors and beneficiaries on the performance of policies, plans and programmes” (2012, p. 17). From this standpoint, the need to develop capacities for evaluation with diverse approaches such as qualitative evaluation, participatory evaluation, gender sensitive evaluation and/or cultural diversity, among others, is observed. This need becomes more relevant since, as will be seen later, although there is a wide range of training offers, they are quite homogeneous in terms of the methods and approaches addressed.

### *Public Perception and Discussion of Evaluation and Evaluation Results*

Evaluation, as an instrument, is mostly known within the community or circle of specialists linked to design, implementation and M&E of programmes, however, outside it, it is not a widely recognised tool.

In Chile, there is currently no discourse or debate on the public agenda about evaluation and its results. However, isolated initiatives are observed, intending to move towards the launching of this debate. These initiatives are generally implemented by the academic community, by specialised study centres and professional networks such as EvalChile.

In the last three years, these initiatives have focused on the framework of international events such as Evaluation Week, promoted by the Centre for Learning in Evaluation and Results for Latin America and the Caribbean and the account for the realisation of a total of eight evaluation dissemination activities, led by civil society organisations, international cooperation projects, evaluator organisations, academia and a public institution during 2017 and 2018. (CLEAR, 2017, 2018).

Evaluation reports issued by the DIPRES and the MDSF in the framework of the M&E systems reviewed in this article are published on the

websites of these institutions. Although there are no data on their use, interviews carried out confirm that these portals are used frequently and that the aforementioned reports are downloaded.

Evaluations commissioned by the Ministries, Undersecretaries and/or Public Services are frequently published on their web pages, however, in those cases in which these reports are not published, there is the possibility for any citizen to request the reports from the agency, and this must be accepted by the institution as mandated by Law 20,285 on Transparency of Public Function and access to information of the State Administration.

Despite the recognition of the practice of publishing the results of evaluations and studies, the VOPE EvalChile argues that there are challenges to the facilitation of access to this data, such as:

- To ensure the publication of all the results of evaluations carried out by public institutions both centrally, regionally and locally.
- To promote the location of data in easily accessible spaces for those who seek them.
- To encourage the publication not only of executive summaries but also of final reports and other documents with a greater level of detail about the process and its results.
- To stimulate the publication of the data used for evaluation.
- To stimulate the publication of results in friendly formats for diverse audiences not necessarily familiar with evaluation.

Neither the contributions of evaluation in general, nor their use, or the results of particular evaluations are present in Chilean media. Although there is no published evidence on the reasons explaining the absence of the practice and findings of evaluation processes in the media, the discourse of those who were interviewed accounts for a limited dissemination of evaluation results. Even though these are public—on many occasions—they are not easily accessible and are not presented in non-specialised readers—friendly formats.

In general, it is observed that the results, recommendations and lessons learned from the practice of evaluation are disseminated and used by a limited group of professionals linked to the design and evaluation of policies and programmes. Neither the evaluation nor its results have been included in the public agenda in Chile and this is an issue that has not

permeated outside academic and specialised consulting circles. Moreover, the full launching of the debates on evaluation in specialised circles is a challenge for those working in positioning them, such as EvalChile.

The discourse of the technical actors consulted does not acknowledge the existence of a demand for evaluations made by civil society or other actors outside the public sector. Maybe there is only a specific and isolated example thereof. At this point, it is important to clarify that there is no significant presence of international cooperation programmes contributing to the adoption of an evaluation practice in Chile, based on a periodic demand for these services, as can be observed in other countries in the region. The demand for evaluation services in Chile is mostly carried out by the State (Olavarría & Peroni, 2012).

However, the analysis of the discourse of the political actors interviewed shows references made to the existence of a demand for the generation of evidence to nurture informed decision-making. This demand would be mainly addressed towards Parliament. And it is described as a sectoral demand emanating from citizen movements that, based upon an analysis of the existing evidence and the demand for new evidence, contribute to the legislative debate. This demand emanates from movements that work in.

Issues that have marked the public agenda in the last four years [...] regarding pensions, childhood matters, human rights matters, linked to indigenous peoples and I would say that issues related with sexual and reproductive rights and also with gender diversity have recently resurfaced. (Interview with Deputy of the Republic)

This tendency to encourage and stimulate an evidence-based public debate, beyond the Budget Law, has been visualised and promoted from a bench of young deputies who have entered the congress in the last period. This poses new challenges to evaluation regarding the need to generate and disseminate evidence on social and political contingency issues.

## PROFESSION/DISCIPLINE

### *Academic Courses of Study, Training, Et Cetera*

There is currently no evaluation chair in Chile, however, there is a wide offer of evaluation training programmes. The academic field addresses evaluation, both at undergraduate and postgraduate level. In the case of

undergraduate studies, economics, administration and sociology schools have evaluation branches, usually between the third and fourth year of professional training.

In the case of postgraduate studies, there is a wide offer that explicitly covers the evaluation of programmes and projects, however, this offer exists almost exclusively at Diploma level. Nevertheless, there are a number of Master's degree programmes having an evaluation training component, such as:

- Master in Public Policy Management, Faculty of Physical and Mathematical Sciences, *Universidad de Chile*
- Master in Public Policy, Faculty of Economics and Business, *Universidad de Chile*
- Master in Management Control, Faculty of Economics and Business, *Universidad de Chile*
- Master in Management and Public Policies, Faculty of Administration and Economics, *Universidad de Santiago de Chile*,
- Master of Public Policy, School of Government, *Universidad Católica de Chile*
- Master in Public Policy, *Universidad de Desarrollo*.
- Among others

Rodriguez Billela presents the training offer in evaluation not only in Chile, but in the region, reporting a significant increase thereof between 2017 and 2018 as seen in Table 6.4 (Rodriguez Billela, 2018).

In comparison with the other countries in the Latin American and Caribbean region, there is a wide training offer regarding evaluation in Chile. However, this offer is quite homogeneous, since it is mostly a diploma offer, not including an offer to obtain a Master's or Doctor's degree. It is mostly rooted in economics schools, it addresses a mostly econometric and almost exclusively quantitative approach in most cases, and it mostly focuses on the economic or social evaluation of projects. Some programmes stand out because of their particularities: it is observed that one of the programmes has a specific focus on territory and community, and one of the programmes is explicitly based on the logical framework methodology, although most of the programmes also implicitly address this methodology and one of them explicitly addresses impact evaluation methodology.

**Table 6.4** Evaluation training programmes (Compiled by authors, based on Rodríguez Billela, 2018)

<i>Programme</i>	<i>University</i>
Diploma in impact evaluation of public programmes and policies	<i>Pontificia Universidad Católica de Chile</i> —Institute of Economy and Poverty Action Lab (JPAL)
Diploma in formulation and evaluation of social projects with a focus on territory and community	<i>Pontificia Universidad Católica de Chile</i> —School of Social Work and Centre for Public Policies UC
Diploma in public policy evaluation	<i>Universidad de Chile</i> —Faculty of Economics and Business
Diploma in preparation, evaluation and management of projects	<i>Universidad de Chile</i> —Faculty of Economics and Business
Diploma in project management and evaluation	<i>Universidad Andres Bello</i> —Engineering School
Diploma in formulation and evaluation of programmes and projects: Logical framework methodology	<i>Universidad Alberto Hurtado</i> —Posgraduate studies school
Diploma in project management and evaluation	<i>Universidad Mayor</i>
Diploma in economic and social evaluation of projects	<i>Universidad de Santiago de Chile</i> —Faculty of Business and Economics
Diploma in project evaluation	<i>Universidad del Desarrollo</i> —Faculty of Economics and Business
Course on evaluation of public policies and programmes	<i>Instituto Latinoamericano y del Caribe de Planificación Económica y Social. ONU</i>

Beyond this stable or permanent training offer, there is an offer of courses structured based on demand and short-term aspects such as competitions and financing funds, specific training needs, alliances between various organisations, among others. In this spirit, there are courses of project evaluation, formulation and evaluation of social projects, educational evaluation—from early childhood to higher education—evaluation of public programmes, and specific evaluation approaches. These courses are mostly implemented by universities and research centres and companies dedicated to capacity development, but also by international organisations, civil society organisations and voluntary organisations of professional evaluators.



### *Professionalisation/Discipline*

In Chile there are currently no professional journals, newsletters or other forms of communication of national scope exclusively dedicated to evaluation. This could be due to the small size of the community of specialists in the field of evaluation in Chile, the poor coordination between actors making up this field and the lack of organisations whose mission aims at strengthening the practice of evaluation and who count on the human and financial resources necessary to carry out these tasks in a sustained manner.

However, there are academic journals on public policies and/or public management that deal with evaluation in some of their numbers. The Public Policy Magazine of the *Universidad de Santiago de Chile* and the Public Policy Studies Magazine of the *Universidad de Chile* illustrate this.

In Chile there is a VOPE called Chilean Evaluation Network: EvalChile. This VOPE has existed since 2009 and is a group of professionals and academics seeking to contribute to the improvement of the quality of public policy evaluations in Chile, through the integrated work of various public, private and civil society actors, in order to promote a culture of evaluation.

The strategic objectives of EvalChile (2018) are:

- To position evaluation in the public space
- To contribute to the quality of evaluations
- To contribute to the generation of a culture of evaluation in public policies
- To contribute to the improvement of public policies
- To contribute to social change.

Although this VOPE has not developed guiding or other similar principles to frame the performance of the tasks of evaluators in the country, during 2017, the EvalChile team actively promoted the dissemination and debate on the Evaluation Standards for Latin America (Rodríguez Billela et al., 2016), as a proposal of guiding principles for evaluative practice in Chile.

These standards aim to “contribute to the development of a common frame of reference that serves as guidance for obtaining evaluations with high quality standards, training and professional practice, facilitating communication between the participating actors, learning and knowledge

generation through professional practice, and the promotion of a culture of evaluation and social responsibility” (Rodríguez Billela et al., 2016, p. 2).

The work of this VOPE is mainly carried out in three courses of action: dissemination, capacity development and articulation among various actors, aimed at strengthening the culture and practice of evaluation in Chile. Regarding dissemination spaces for the strengthening of evaluation culture, discussion panels and working groups on evaluative research are held in academic congresses. As for capacity development, training courses on issues related to M&E with different approaches are carried out. Concerning the articulation of actors to strengthen the culture of evaluation, projects are designed and implemented in collaboration with other relevant actors at national and international level (EvalChile, 2018). For example, as indicated in the interviews conducted, since 2018 the VOPE has been working on two projects supported by the international community of evaluators, a Flagship Project—financed by EvalPartners—and a Peer to Peer project developed in conjunction with the Paraguayan Evaluation Network and funded by the American Evaluation Association. Both projects aim to strengthen the involvement of civil society actors in the M&E of the Sustainable Development Goals and in the voluntary national reports on this topic.

Both the academic community, the public institutionality in charge of administering the M&E Systems and the Chilean Evaluation Network acknowledge that the evaluation offer in the Chilean market is formed mainly by evaluators who have developed their expertise in this field from professional practice, rather than having academic certifications as professional evaluators, nonexistent at the national level nowadays. These actors recognise to what extent the systematic implementation of outsourced evaluations since the 1990s has allowed the consolidation of a wide market of qualified evaluators, who work independently, in universities and study centres or in consulting companies.

In its analysis of M&E systems in Chile, the World Bank comes, among others, to the conclusion that this system has introduced mechanisms to promote the objectivity, impartiality and reliability of evaluation processes and findings; and that the involvement of external experts in the design and implementation of evaluation processes has resulted in an increase of the expertise with which the evaluations are implemented, as well as their validity and reliability (Dussange, 2011).

Those who work in evaluation come mostly from the fields of economics, engineering and sociology, however, the presence of professionals from various disciplines is observed. These people perform a variety of tasks related to the evaluation process (management, design, data collection and analysis, communication of results, among others) and in various areas of evaluation.

A large number of those who recognise themselves as experts in evaluation in Chile divide their professional practice between academia and consulting. This figure provides flexibility to serve as independent evaluators, being part of the DIPRES and MDSF expert registries, among others, and joining evaluation teams associated with consulting companies, universities and/or study centres if they are required to respond to requirements of greater complexity or intensity.

There are no certification systems for evaluators in Chile and there is also no technical authority in charge of conciliation in case of dispute, it is up to the Courts of Justice to settle any disputes regarding service delivery contractual aspects. However, a public registry of evaluators has just been enabled, in the DIPRES, constituting a first database to be used by the Ministry of Finance.

### *Compliance with Quality Standards and Obligations*

The quality standards of an evaluation, in terms of design, implementation and dissemination of its results are delimited and explained in its terms of reference. As previously noted, most evaluations carried out in Chile are outsourced and this process is carried out through a public procurement system or its equivalents in the private sphere, in which a series of technical and economic offers compete. The initial eligibility criterion thereof is the degree to which the proposals are relevant regarding the requirement reflected in the terms of reference.

As a general trend, the terms of reference detail the approaches, methodologies and even (sometimes) the instruments that should be used in an evaluation. Adjusting these definitions requires a negotiation and consensus process with the institutional technical counterpart.

In Chile, contracts concerning the provision of evaluation services usually incorporate fines in the event of non-compliance with the various contractual conditions. Due to this condition, it is a widespread practice to coordinate and jointly address—both the external evaluation teams and

the requesting institution—those difficulties putting at risk the implementation of the evaluation according to the quality standards and the design initially conceived.

## CONCLUSION

M&E are fully installed in the endeavour of the Chilean State, throughout the evaluation endeavour of the DIPRES and the MDSF. The tools, instruments and procedures associated with these M&E systems are currently institutionalised in the tasks carried out by the different Ministries and services.

This institutionalisation of M&E systems is mainly due to the fact that the Management Control System has been operating for decades, led by the DIPRES, and to the accumulated experience the MDSF has acquired through the application of monitoring and evaluation instruments. Chile is a pioneer in the use of M&E systems in the framework of the budget cycle in the region and it has fully installed a budgeting for results approach, which ranges from systems and platforms to an organisational culture in the different services delivered.

The information generated by M&E systems is mainly used for budgetary decision-making, in terms of public spending and investment. Moreover, it is used—although to a lesser extent—for improvement regarding the design and implementation of programmes, as of commitments acquired among the institutions evaluated and those who implement M&E Systems. While this use of evaluation results aiming to provide feedback to programmes and services in order to improve their practices is important, there are challenges to the increase of the use of evaluation for improving policies, plans and programmes, for the rendering of accounts, for planning and for the stimulation of citizen participation in the evaluation processes.

M&E systems work almost exclusively based on the outsourcing of evaluation services, which has allowed the development of a competitive market of experts in evaluation, trained in professional practice, without professional certifications delivered nationwide. This market of highly competent evaluators has emerged in response to the demand in the last two decades and it is composed of evaluators who carry out their work independently, teams linked to universities and study centres and teams of specialised consulting companies.

The evaluation training offer is broad, compared to other countries in the region, however, this offer is quite homogeneous and linked to the mainstream evaluation exercise in Chile: quantitative, related to the economic and social evaluation of projects and based on the logical framework methodology, with few exceptions.

The incorporation of civil society in the evaluation processes is scarce, however, there are exceptions concerning the State, as well as some limited examples of evaluations carried out by civil society organisations themselves. Strengthening this incorporation in an effective, efficient and relevant manner and strengthening evaluation in civil society organisations faces not only political and managerial but also technical and methodological challenges. It should be remembered that many social policies are implemented by civil society organisations, therefore their role as co-implementers, can be detrimental to the exercise of independent evaluation (accountability). It is possible that the absence of this exercise may have an impact on the absence of new perspectives from which to assess evaluation (e.g. evaluation from a gender perspective, from the perspective of Human Rights, etc.)

The debate on evaluation in Chile is in the process of being launched in a specialised circle of academics, consultants and researchers, however, this launching requires strategies of encouragement and articulation of public and private actors. However, it is not a topic on the agenda outside these circles.

The strengthening of the use of evaluation and its assessment as a public good, the development of capacities in the various actors that make up the (eco) evaluation system, the valuation of diverse approaches and methodologies, the involvement of civil society and the strengthening of a community of evaluators are some of the main challenges that arise from this analysis in order to close a virtuous circle between demand, implementation and use of evaluation in the Chilean state.

### LIST OF ABBREVIATIONS

BIP	Integrated Bank of Projects
BIPS	Integrated Bank of Social Programmes
CLAD	Latin American Centre for Development Administration
DIPRES	Budget Office
EFA	Focalized Scope Evaluation
EGI	Evaluation of Institutional Expenditure

EI	Impact Evaluation
EPG	Evaluation of Government Programmes
EvalChile	Chilean Evaluation Network
MDSF	Ministry of Social Development and Family
M&E	Monitoring and Evaluation
ODEPLAN	National Planning Office
SEIA	Environmental Impact Assessment System
VOPE	Voluntary Organisation for Professional Evaluators

## REFERENCES

- BIPS. (2018). *Budget of social programmes BIPS*. <https://programassociales.ministeriodesarrollosocial.gob.cl/programas>. Accessed on 2 August 2020.
- CLEAR. (2017). *Semana de la evaluación en América Latina y el Caribe 2017*, Proceedings. Centro de Investigación y Docencia Económica, Centro para el Aprendizaje en Evaluación y Resultados para América Latina y el Caribe. <https://clear-lac.org/wp-content/uploads/2019/09/18.-Memorias-EVAL2017.pdf>. Accessed on 11 August 2019.
- CLEAR. (2018). *Semana de la evaluación en América Latina y el Caribe 2018*, Proceedings. Centro de Investigación y Docencia Económica, Centro para el Aprendizaje en Evaluación y Resultados para América Latina y el Caribe. [https://clear-lac.org/wp-content/uploads/2019/09/19.-Memorias-EVAL2018\\_web.pdf](https://clear-lac.org/wp-content/uploads/2019/09/19.-Memorias-EVAL2018_web.pdf). Accessed 10 August 2019.
- Darville, P., Díaz, R., & Leiva, J. (2017). *Cobertura evaluación de programas e instituciones públicas*. Departamento de Evaluación. División de Control de Gestión Pública. Ministerio de Hacienda. Gobierno de Chile.
- Decreto Ley Número 1.263 Orgánico de la administración financiera del Estado. <https://www.bcn.cl/leychile/navegar?idNorma=6536&idVersion=2020-04-02&idParte=8746724>. Accessed on 15 August 2019.
- DIPRES. (2019). *Evaluation results*. Second Quarter 2019. Ministerio de Hacienda. Gobierno de Chile.
- Dussange, M. (2011). *Chile's monitoring and evaluation system, 1994–2010 (English)*. PREM notes; no. 14. Special series on the nuts and bolts of m&e systems. World Bank. <http://documents.worldbank.org/curated/en/740771468225872299/Chiles-monitoring-and-evaluation-system-1994-2010>. Accessed on 5 August 2019.
- EvalChile. (2018). *Informe de rendición de cuentas 2017–2018*. Internal document. Santiago de Chile.
- Irrazaval, I., & De los Ríos, B. (2015). Una construcción dinámica: El sistema de monitoreo y evaluación de Chile. In G. Pérez Yarahuan & C. Maldonado

- Trujillo (Eds.), *Panorama de los sistemas nacionales de monitoreo y evaluación en América Latina* (pp. 113–148). Centro de Investigación y Docencia Económicas-Centro CLEAR para América Latina. México.
- Ley 20,285. 11-August 2008. *Ministerio Secretaría General de la Presidencia—Ley Chile—Biblioteca del Congreso Nacional*. <https://www.bcn.cl/leychile/navegar?idNorma=276363>
- Ley N° 20530. 13-October 2011. *Ministerio de Planificación—Ley Chile—Biblioteca del Congreso Nacional*. <https://www.bcn.cl/leychile/navegar?idNorma=1030861&idParte=9193568&idVersion=2019-04-16>. Accessed on 10 August 2019.
- Ministerio de Desarrollo Social y Familia. (2019). [www.programassociales.cl/2019](http://www.programassociales.cl/2019). Accessed on 14 August 2019.
- MDSF. (2019). *Consejo de la Sociedad Civil*. <http://participacionciudadana.ministeriodesarrollosocial.gob.cl/consejo-de-la-sociedad-civil/consejo-de-la-sociedad-civil>. Accessed on 11 October 2019.
- OECD. (2012). *La evaluación de leyes y regulaciones. El caso de la cámara de diputados de Chile*. [https://www.oecd-ilibrary.org/governance/la-evaluacion-de-leyes-y-regulaciones\\_9789264176362-es](https://www.oecd-ilibrary.org/governance/la-evaluacion-de-leyes-y-regulaciones_9789264176362-es). OECD Publishing. Accessed on 7 August 2019.
- Olavarria, C., & Peroni, A. (2012). La evaluación en el Estado Chileno, Avances y Desafíos. *Revista Políticas Públicas*, 4(2), 5–20, Universidad de Santiago de Chile.
- Ríos Hess, S. (2007). *Fortalecimiento de los sistemas de monitoreo y evaluación en América Latina, diagnóstico de los sistemas de monitoreo y evaluación en Chile*. CLAD-BID.
- Rodríguez Billela, P. (2018). *Más de cuarenta posgrados en evaluación en América Latina y el Caribe*. Al borde del caos. <https://albordedelcaos.com/2018/05/17/mas-de-cuarenta-posgrados-en-evaluacion-en-america-latina-y-el-caribe/>. Accessed on 15 August 2019.
- Rodríguez-Billela, P., Valencia, S. M., Alvarez, L. S., Klier, S. D., Hernández, A. L. G., & Tapella, E. (2016). *Estándares de evaluación para América Latina y el Caribe*. Ciudad Autónoma de Buenos Aires.
- Sánchez, C., & Miranda, D. (2019). *Experiencias de evaluación participativa en intervenciones del programa Servicio País (Chile). Incorporando a la comunidad en la evaluación de programas sociales*. <https://evalparticipativa.files.wordpress.com/2019/05/piloto-evaluaciones-participativas-en-servicio-pac3ads-tapa.pdf>. Accessed on 12 August 2019.
- Servicio de Evaluación Ambiental. (2019). *¿Qué es SEIA?* <https://www.sea.gob.cl/sea/que-es-seia>. Accessed on 7 August 2019.
- TECHO. (2019). *Informe de evaluación de la mesa de trabajo del campamento Santa Teresa*. <https://evalparticipativa.files.wordpress.com/2019/05/informe-de-evaluacion-de-la-mesa-de-trabajo-del-campamento-de-santa-teresa.pdf>. Accessed on 8 August 2019.



## Evaluation in Colombia

*Maria Gladys Álvarez Basabe,  
María Elena Manjarrés De Mendoza,  
Luz Stella Uricoechea Morales,  
María Paula Uricoechea Castellanos,  
and Belkis Esperanza Vergara Pérez*

### GENERAL OVERVIEW

The Republic of Colombia is located in South America. It is the only nation having coasts in both the Pacific and the Atlantic Ocean and it is the second most biodiverse country in the world.<sup>1</sup> Its multicultural and multiracial population encompasses over 48,835,324 inhabitants (DANE, 2018), organised in 32 departments and the Capital District of Bogotá, the seat of government. It ranks as the fourth largest economy in Latin

---

<sup>1</sup> Retrieved from: <https://www.minambiente.gov.co/index.php/noticias-minambiente/4317-colombia-el-segundo-pais-mas-biodiverso-del-mundo-celebra-el-dia-mundial-de-la-biodiversidad>. Last access: June, 2020.

<sup>2</sup> Retrieved from: <https://www.colombia.com/colombia-info/informacion-general/>. Last access: March 15, 2020.

---

M. G. Álvarez Basabe (✉) · M. E. Manjarrés De Mendoza ·  
L. S. Uricoechea Morales · M. P. Uricoechea Castellanos · B. E. Vergara Pérez  
Instakids S.A.S., de Bogotá, Colombia



America, with an annual growth of 3.7%. Its Gross Domestic Product (GDP) also ranks fourth in the region and 30th worldwide.<sup>2</sup> Colombia is a State of unitary, social, democratic and presidential law, with autonomy from its territorial entities. The branches of public power are: legislative, executive<sup>3</sup> and judicial.

## INSTITUTIONAL STRUCTURES AND PROCESSES

### *Evaluation Regulations*

The institutionalisation of evaluation in Colombia is a purpose under continuous construction. It seeks to rationalise and make public action more effective, as well as to improve the efficiency and effectiveness of the State, following both the New Public Management (NGP) and the *Gestión por Resultados* (Results-Oriented Management approaches, GPOR), implemented in the country since the nineties (Serra et al., 2007, p. 7). The last three decades have been dynamic concerning the issuance of evaluation regulations, aimed at strengthening the institutional framework and consolidating the practices of monitoring and evaluation of programmes and public policies in the country. They “have intended to establish a culture of evaluation that allows ...to increase transparency and, above all, provide elements in order to improve public action and inform citizens and public institutions of its effectiveness and efficiency” (Roth, 2009, p. 1).

The entities supporting the President, as the maximum guiding instance of national planning, are the *Consejo Nacional de Política Económica* (National Council for Economic and Social Policy, CONPES) and the *Departamento Nacional de Planeación* (National Planning Department, DNP). In the 1980s, with the issuance of Decree 3152 of 1989 and extraordinary Decree 2410 of 1989 (Art. 1–3), the organic structure of the DNP was modified and its functions were determined, as the main institutional body of the administration in charge of formulating and elaborating the general plans and programmes of economic

<sup>3</sup> At the national level, the Executive Branch is composed of the President of the Republic, cabinet ministers and the directors of administrative departments. At departmental and municipal level, it consists of governments and mayors (...) as well as public establishments and state industrial or commercial companies, as defined in the Political Constitution of 1991. (Art. 1, 113, 115).

and social development. In addition, it coordinates the National System for Planning and Evaluation of Public Policies.<sup>4</sup>

The Political Constitution (1991, Art. 343, 344) grants the DNP the function of designing and organising the evaluation systems of management and results achieved by public investment in the country. Territorial entities were granted the responsibility of planning and evaluating the management and results of departmental, district and municipal public policies. In both cases, the democratic mechanisms of social control and surveillance are defined (Art. 103). Through the approach of a GPOR, the State-citizenship relationship was deepened (CONPES 2790, 1995, p. 1).

The following two decades were significant regarding the issuance of the regulatory legal framework<sup>5</sup> required for creating and consolidating the conceptual and operational capacity of government agencies, including the DNP and its dependencies. With CONPES 3294 (2004, p. 1), efforts were directed at institutionalising evaluation, monitoring, and accountability of government policies and programmes as a permanent practice of the State.

The Monitoring, Follow-up, Control and Evaluation System (SMSCE) is created and regulated<sup>6</sup> in order to ensure the efficient and effective use of the resources coming from the General Royalty System (SGR), in order to strengthen transparency, citizen participation and Good government. Its efforts were oriented to the construction of an evaluation vision for the SGR, so that each one of its components (Monitoring, Follow-up, Control and Evaluation) were conceptually and operationally differentiated and their scope defined.<sup>7</sup> The National Management and Results

<sup>4</sup> They are made up of: the Ministries' planning offices, administrative departments, regional planning councils, planning administrative offices or departments, planning councils and finally, the entities attached to the DNP.

<sup>5</sup> The following regulations are issued to restructure the DNP: Law 152 1994 Decree 2167 of 1992 and Law 489 of 1998 - Decree 3517 of 2009, Decree 1832 of 2012, Decree 1163 of 2013 and Decree 2189 of 2017.

<sup>6</sup> The SMSCE is created in Sub section 7.1 of paragraph 3 of article 361 of the Political Constitution and it is regulated in Decree 414 of 2013, Law 1530 of 2012, Decree 817 of 2014, Decree 1083 of 2015 and Decree 2189 of 2017. Its administration is assigned to the DNP Bureau for the Surveillance of Royalties.

<sup>7</sup> For greater control of SGR resources, a technical audit, the value of which cannot exceed 10% of the total project, is hired in order to complement the evaluation.

Evaluation System (SINERGIA) is created and regulated<sup>8</sup> in order to improve transparency in public management and to introduce a culture of evaluation within the entities involved, as well as to strengthen capacities concerning the management of public investment. In addition, it assumes the function of establishing guidelines, instruments and mechanisms for monitoring and evaluating the results and impacts of programmes and projects implemented by governmental, national and territorial institutions. Both systems actively incorporate the control of citizens and organised civil society.

Simultaneously, the State implements procedures and methodologies for evaluating public administration management. In 1998, the *Modelo Integrado de Planeación y Gestión* (Integrated Planning and Management Model, MIPG)<sup>9</sup> was adopted to direct, plan, implement, monitor, evaluate and control its management. The Model integrates other public administration systems, such as internal control and quality management. In 2017, territorial entities must adopt the updated version of the MIPG, which is the “current public management system, (...) that allows its entities to orient themselves towards the fulfillment of its objectives” (Hernández, 2018, p. 447). “The primary goal of these reforms was to establish instruments that would allow Colombia to analyse its state capacity to articulate efficient solutions to the country’s economic and social problems, through the permanent evaluation of the strategies and results deriving from the different programmes implemented within the framework of public policies included in the National Development Plan (PND)” (DNP, 2010, p. 14).

### *Evaluation Practice*

In the last three decades, due to the demand for tools concerning the monitoring and evaluation of public programmes and policies (Roth,

<sup>8</sup> SINERGIA is created in CONPES 2688 Resolution 63 of 1994 (Art. 1,2,4) and is regulated in: Decree 2167 of 1992 Art. 22, Decree 1832 of 2012 art 3 and Decree 2189 of 2017, (Art. 3.15). Its administration is assigned to the DSEPP Initially called Special Division of Evaluation and Management Control (DEE).

<sup>9</sup> Law 489 1998, Art. 15, Law 1753 of 2015 Art. 133 and Decree 1419 of 2017. The MIPG is administered by the Departamento Administrativo de la Función Pública (Administrative Department of the Public Function, DAFFP).

2009, p. 1), the country consolidated the legal, institutional and operational framework of SINERGIA, SMSCE and the MIPG, recognising that this is not a finished task, because due to the dynamics of public management, these are living systems in constant transformation, adaptation and redesign (Pérez Yarahuán & Maldonado Trujillo, 2015, p. 18).

As a result, public entities nowadays rely on the administrative structure and technical capacity to monitor and evaluate strategic interventions, management and institutional performance, which was “achieved through expert advice and courses on monitoring and evaluation techniques” (DNP, 2019, p. 7). Likewise, recognising institutional and territorial autonomy and its particularities, the roles of each of the actors of the National Planning and Evaluation System and their scope were defined, so that they carried out their evaluations, through interaction with multiple actors with different interests, mandates, incentives and different roles and from different perspectives (Pérez Yarahuán & Maldonado Trujillo, 2015).

In this context, evaluation has been present in the public administration for more than three decades, counting on the resources required for its implementation. The SMSCE is financed with an annual percentage of the SGR resources.<sup>10</sup> Evaluations carried out by MIPG and SINERGIA are financed with the nation’s own resources (DNP, 2019). According to the *Dirección de Seguimiento y Evaluación de Políticas Públicas* (Bureau for Monitoring and Evaluation of Public Policies, DSEPP), there are financial and human restrictions for the development of evaluations in this entity (DNP, 2019), therefore, the number of projects evaluated each year is not significant compared to those implemented and the number of projects evaluated annually is fluctuating. In 2003 one project was financed, while in 2010 the total amount was 35 (Table 7.1). Between 1997 and 2007, 45 evaluations were carried out and from 2008–2018, 221, with a 390% increase. As of 2018, the number of evaluations per year has decreased, to the point that in 2019 only three evaluations were carried out.

Apparently, in Colombia, the evaluations of the policies, programmes and projects derived from the PND are financed. Consistent with this, it was found that 86% of evaluations have been financed with public money, 73% of which were covered with investment resources from the General

<sup>10</sup> Clause 2 of Legislative Act 5 of 2011.

**Table 7.1** Evaluations carried out per year (SINERGIA Repository of evaluations) (own development)

<i>Year</i>	<i>Number of evaluations</i>	<i>%</i>	<i>Year</i>	<i>Number of evaluations</i>	<i>%</i>
1997	1	0,4	2010	35	12,9
2001	1	0,4	2011	25	9,2
2002	2	0,7	2012	33	12,2
2003	1	0,4	2013	18	6,6
2004	4	1,5	2014	12	4,4
2005	9	3,7	2015	10	3,7
2006	12	4,4	2016	11	4,1
2007	15	5,5	2017	20	7,4
2008	30	11,1	2018	12	4,4
2009	15	5,5	2019	2	0,4
			2020	3	1,1
TOTAL: 271					100

Budget of the Nation and 13% from the SGR. Only a 14% was financed with International Cooperation resources. This explains why 83% of the evaluations reviewed correspond to public policies, 33% to programmes and 17% to PND projects carried out in the final phase of implementation, thus prioritising interventions for those subject to reformulation (DNP, 2019). The evaluation of other projects is financed with external debt<sup>11</sup> and International Cooperation resources. For this purpose, according to the interviewees, entities create their own evaluation models, methodologies and instruments, as well as they define the frequency of evaluation and the modality of State participation in these processes.

Additionally, evaluation culture, transparency in public management and citizen control are actively encouraged, through the use of solid web tools that provide feedback to public policy (DNP, 2015). With this, progress has been made in the appropriation and use of Information and Communication Technologies (ICT) in public administration.

<sup>11</sup> The Inter-American Development Bank (IDB) uses the ‘Development Effectiveness Framework’ tool and the World Bank (WB), relies on the Independent Evaluation Group (IEG).

Hence, the SMSCE supports its actions on the GESPROY–SGR<sup>12</sup> platform, which is transversal to the project execution cycle, and provides a detailed overview of its progress.<sup>13</sup> With its support, 4,005 projects have been followed up, corresponding to (28%) of the 14,563 projects approved by the SGR.<sup>14</sup> The MIPG has a microsite<sup>15</sup> to collect information,<sup>16</sup> carry out self-diagnosis, assess institutional performance and compare the entities, with differential criteria and considering the specificities of national and local politics (DAFP, 2017). At the same time, entities build their own management tools tailored to their purpose. That is the case of the Integrated Management System (GIS) of the Ministry of Trade, Industry and Tourism, which, according to the interview carried out, ‘allows Strategic Sector Planning (PES), identification of progress made towards the goals, evaluation of performance and institutional management regarding economic and social development of the country’.

SINERGIA supports management on an online platform made up of two modules<sup>17</sup>: evaluation of the results and impacts of public policies, in compliance with the country’s development objectives and monitoring of government goals<sup>18</sup> to continuously monitor the performance of public entities as an input for taking corrective actions to achieve the objectives of the PND. The analysis of the SINERGIA evaluation module allows us to conclude that, since 1997, 271 project evaluations have been carried out, financed with both investment and SGR resources. These are

<sup>12</sup> GESPROY is created in Decree 414 of 2013, Circular of DNP 62 of 2013, Decree 817 2014 Article 68.

<sup>13</sup> Retrieved from: <https://www.sgr.gov.co/SMSCE/EvaluaciónSGR.aspx>. Last access: March 20, 2019.

<sup>14</sup> Retrieved from <https://www.sgr.gov.co/Proyectos/BancodeProyectos.aspx>. Last access: March 20, 2019.

<sup>15</sup> Retrieved from: <https://www.funcionpublica.gov.co/web/mipg/resultados-medicion>. Last access: June 23, 2020.

<sup>16</sup> Information is recorded in the Single Management Progress Report Form.

<sup>17</sup> Retrieved from: <https://sinergiapp.dnp.gov.co/#HomeSeguimiento>. Last access: April 10, 2020.

<sup>18</sup> Until 2010 it was called SIGOB (Management and Monitoring of Government Goals System) and from 2011, SISMEG (Monitoring System for Government Goals).

selected annually according to strategic criteria defined in the methodology proposed by the DSEPP,<sup>19</sup> among them: strategic interventions, feasibility of the evaluation to guarantee the development and subsequent use of the results, volume of investment and beneficiaries served (DNP, 2019). 53.8% of evaluations are internal and 46.2% external; while in the sample of 30 evaluations reviewed,<sup>20</sup> the relationship is 80% internal and 20% external. According to the SINGERGIA flowchart, external consultants are the ones who carry out the evaluations, while the entity's team coordinates, designs and supervises the evaluation and facilitates the use of evidence (DNP, 2019).

External evaluations were carried out by consultants with a recognised track record on the subject, as evidenced in number three of this article. These actors propose recommendations to the implementing entity for the most effective and efficient use of resources, as well as they take corrective measures to improve State action. 35.8% of these evaluations are results evaluations, 26.2% impact evaluations, 11% institutional and operations evaluations (Fig. 7.1). In both samples, it is found that there are few or no experimental, cost/benefit or cost/effectiveness evaluations that could measure the direct impact of public policy improvement.

The DSEPP selects the interventions to be evaluated and includes the largest number of public administration sectors (DNP, 2019); 20 out of the 24 sectors classified by SINERGIA rely on evaluations, the highest 'development' score being found in health and social protection 6%, the Presidency 11%, social inclusion and reconciliation 10%, *Ciencia, Tecnología e Innovación* (Science, technology and innovation, CTelI) 9%. The four sectors with the lowest 'development' scores are: Treasury and Public Credit, Sport and Intelligence and Foreign Relations, which to date have not carried out any evaluation (Fig. 7.2).

The interviews carried out report that there is no specific evaluation regulation for each of these sectors. The country has mandatory regulations in all state entities to guarantee their effective control. However, some of these entities have defined a methodological and legal framework to evaluate their objectives. One case is the Ministry of Information Technology and Communications (MINTIC), which, in compliance with

<sup>19</sup> Every year, the Annual Evaluation Agenda is made up of strategic public interventions, according to Art. 13–15 of Decree 414 of 2013.

<sup>20</sup> They correspond to the policies, programmes and projects of the PND implemented by the Presidency of the Republic.

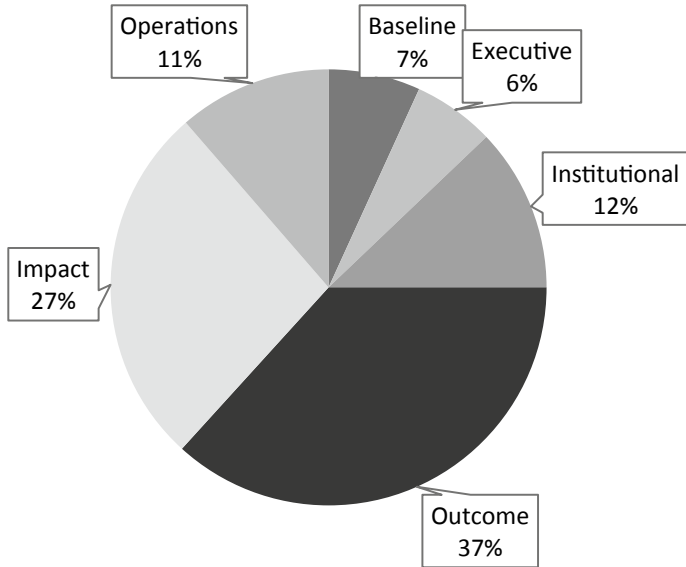


Fig. 7.1 Typology of evaluations (SINERGIA repository)

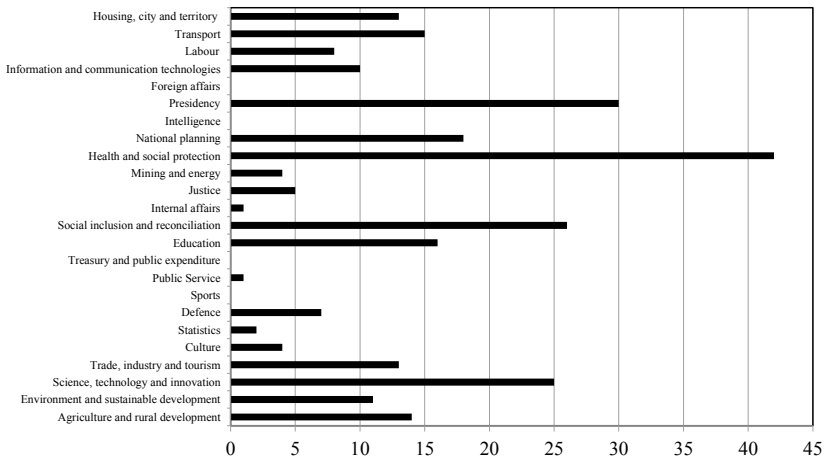


Fig. 7.2 Sectors carrying out evaluations (SINERGIA repository)



ministerial function No. 15 of article 1 of Decree 1414 of 2017, defined the process of Monitoring and Evaluation of ICT Policies.

### *Use of Evaluations*

In the last three decades, the importance of promoting activities and strategies for the dissemination, use and appropriation of evaluation results by different actors in society has been emphasised. SINERGIA assumes a leading role in this task, which is recognised by the government sources consulted. They point out that the most widely used means of dissemination is its Web platform, in whose repository you can consult the topics that have been or are being evaluated, the reports evaluations, advances in programmes and policies derived from the PND. However, civil society members surveyed did not report the use of SYNERGY.

Professional groups that regularly use assessment results are still mainly programme and policy managers (76%); less frequently, decision-makers (7%). According to the government actors interviewed, the participation of citizens and civil society organisations is still very incipient, which is expressed more clearly in chapter two of this document. This result coincides with the fact that, in 50% of the reports reviewed, evaluations are used for decision-making and in 23%, they serve the purpose of reorienting PND projects and programmes.

Likewise, most parliamentarians among which a survey was carried out affirm that parliamentary structures sometimes use the results of evaluations for their political work, especially for the design of bills. 83.3% of evaluation results are used to support political control debates and bills, 50% to justify their presentations in the Congress of the Republic and 16.7% for decision-making and accountability. In the open-ended questions, they express that “most of the time, in order to support the political debate and the bills, reports from the control bodies such as the Office of the Comptroller General and the Attorney General of the nation are used, as well as studies of academic or research institutions” (Appendix 7.1).

Multilateral organisations, such as the World Bank (WB), use evaluation results to formulate policies that require reliable information and to make decisions regarding debt. For the *Agencia Presidencia de Cooperación Internacional* (Presidency Agency for International Cooperation, APC), evaluation results are used as an input for the negotiation of new country strategies and mixed cooperation commissions, project proposals,

territorial delimitation for future strategies, programmes and projects, among others.

However, the extensiveness of the public policy evaluation concept is evident and, therefore, in their responses it is associated with studies made by academic or research institutions, such as internal control audits and external control bodies concerning the General Comptroller of the Nation and the Office of the Attorney General, and, finally, Accountability to citizens and performance evaluation of the work team. For example, the Ministry of National Education, does so in terms of Inspection and Surveillance, in accordance with the provisions of Decree 5012 of 2009 and Law 1740 of 2014.

Finally, in order to guarantee the quality of evaluations and the opportunities for improving the information generated by the DNP or by other actors, this entity made the *Archivo Nacional de Datos* (National Data Archive)<sup>21</sup> available to citizens. Through it, users can access the metadata and microdata of the evaluations carried out (DNP, 2019, p. 13). With the same purpose, a guide for the construction and analysis of indicators was designed. It substantiates the strategy of “the value chain” (DNP, 2019, p. 7) and the committees of government entities that guarantee the veracity and quality of the information published on their portals, which is one of the most important assets of public entities.

In this gradual process concerning the institutionalisation of evaluation, significant progress has been made in legal matters, institutional and conceptual strengthening, as well as the introduction of a practice aimed at determining the degree of efficiency and effectiveness of administration and progress in implementation of public policies, in relation with the country goals established in the PND and the consolidation of a culture of evaluation in the country. However, at this time, Colombia faces some challenges in terms of ensuring that all its entities implement their regulated monitoring and evaluation policies and strategies, as well as in terms of measuring their impact concerning the real impact of the State’s action on the problems it seeks to solve, beyond attesting that the proposed activities were carried out to the satisfaction of the ‘value chain’ of each project or programme.

<sup>21</sup> [www.anda.dnp.gov.co](http://www.anda.dnp.gov.co)

## DISSEMINATION IN SOCIETY AND ADAPTION (SOCIAL SYSTEM)

Analysing the institutionalisation processes of evaluation in Colombia, from the point of view of civil society participation, involves addressing aspects such as the legal structure that supports social participation in state actions, the practices that civil society develops regarding evaluation and the participation culture of civil society.

Legal structure supporting the participation of civil society in the evaluation processes concerning projects, policies and programmes.

Citizen participation is reflected in people's consent to be part of the decisions affecting them, to exercise control over public authorities and management and to define the parameters regarding its implementation, especially procedural ones. In this sense, Colombian civil society plays a fundamental role in the public policy cycle, both as the main user and client of the goods and services offered by the Government and as the observer of public resources.

The main laws and regulations that establish and support participation in Colombia are presented below:

The Political Constitution of 1991 granted all citizens the opportunity to actively participate and intervene in the control of public administration. At the same time, it determined the way in which citizens participate in the planning, monitoring and surveillance of the results of state management.

Law 134 of 1994 is consolidated under three fundamental pillars: (a) the mechanisms for citizen participation and both the flexibility this law promotes and guarantees in order to use them frequently and effectively; (b) Public Accountability and Public Social Control, and (c) the coordination and promotion of citizen participation, through the National Participation Council as a constituent element of the National Participation System.

Law No. 850 of 2003, on Citizen Oversight, "establishes a mechanism that allows citizens or different community organisations to exercise vigilance over public management in those areas, aspects and levels in which public resources are fully or partially used, and with respect to the administrative, political, judicial, electoral, legislative and control bodies".

Decree No. 695 of 2003 "determines the objectives and functions of the Participation and Strengthening of Democracy Fund".

Law No., 1257 of 2008, “which establishes norms of awareness, prevention and punishment of forms of violence and discrimination against women”.

Statutory Law 1757 of 2015 “by which provisions are made regarding the promotion and protection of the right to democratic participation”. Its purpose is to promote, protect and guarantee modalities of the right to participate in political, administrative, economic, social and cultural life, and also to control political power.

This legislation framework allows us to understand the way in which citizen participation is supported by state actions, as well as the way in which it regulates civil society participation in the design, implementation and results of the evaluation of programmes and public policies, either with State investment or financing from other national and international agencies.

### *Institutionalised Use of Evaluation by Civil Society*

The usual practice in the country concerning the use of evaluations, to provide knowledge and processes through referendums or community-based political decisions, is aimed at expanding knowledge in different sectors of the country and supporting decision-making in civil society or at political levels. This practice seems not to be fully institutionalised: 75% of evaluations and their results are used for political decision-making or to support referendums and new programmes or projects of communities and social sectors. It is important to note that since 1991, the space for community participation has been expanded and the use of this practice is growing, as noted by the government actors interviewed and reported in the first part of this document.

Surveys carried out show that the frequency of evaluation use represents 75% of civil society, however, 25% do not use them due to obstacles such as: (a) poor training and schooling in communities, particularly in the popular sectors, (b) political interests that impede community participation, or d) poor disclosure of the results of the evaluations carried out. The results are used, by three quarters of respondents, mainly individual citizens, civil society organisations and private companies; 65.55% to support community interests, 50% to design a project or social programme, 12.5% to monitor economic resources and their compliance. In the same percentage, they are used to support academic

studies. Regarding the frequency of its use, the percentage reaches 12% (yes or no) without reporting the reasons (Appendix 7.2).

Participation in the evaluation processes varies: 50% of respondents have participated as interviewees; 12.5% as users of programmes or projects; 2.5% as interviewers and 12.5% because they are interested in the topic. The frequency of participation concerning organisations or individuals is the following: 37.5%; once a year; 10% twice a year; 2.5%; according to the project cycles; 2.55% because of political control; 12.5% because of the topic and its specificities; and 2.5% when it is required by the state. In general, there are no obstacles of any kind for participation in these processes (Appendix 7.3).

In particular, there is very little use of the results of public programmes and policies evaluations, in which representatives of civil society participate or from which they benefit. It is worth noting that those who use the evaluation and its results, very infrequently, are the participants of social organisations and foundations, Non-profit Non-Governmental Organisations (NGOs) and those who use them the most are political sectors, especially parliamentarians, at national or regional level, who take advantage of experience regarding the development of projects or keep them as an information base to propose new projects, develop referendums and argue the need for new investments and donations from state or national and international agencies. Regarding this discussion, it is considered “that the theory and, even more so, the practice of evaluation, are relatively recent in the field of development actions. Although its beginnings can be located in the 1950s, its subsequent evolution has been slow and uneven” (Cohen & Franco, 1998, p. 1).

### *Public Perception and Discussion of Evaluation and Its Results*

Regarding the knowledge that members of civil society have about the evaluation instruments to be applied, for 12.5% of respondents these are known, for 50%, little known, for 37.6%, they are unknown. Regarding the availability of evaluation reports for citizens, 87.59% of them are available; and 50% are publicly discussed. It should be noted that, generally, State institutions present their Accountability annually and interested citizens can attend the event. The discussion is frequently attended by 50% of the participants, whose motivation is fostered through the media.

12.5% of evaluation reports are available to the public in written media; 75% in portals of public and private institutions; 37.5% in archives of

the nation, region or municipalities and 12.5% are not available. These results are explained by the fact that very few participants know the instruments used for the evaluation of projects and programmes, or they have never known these instruments. The opposite occurs when a large majority has a willingness to read and review the reports made available to the public, as indicated above, through institutional portals and, less frequently, through previously related print media.

Despite this availability of reports, it seems that members of social organisations, on their own initiative, are not interested in reading or discussing them. They participate when there is insistence on the part of the foundations, NGOs or the state offices, so that they are known and discussed. In this same sense, Cohen and Franco (1998, p. 71) add that an important point of evaluation and its results is that “(...) it must analyse the real validity of the objectives declared in the project documents, to appreciate if they match the truly persecuted goals; a project must have intended, expected, positive and relevant effects”, and perhaps social groups are not very interested in learning about the evaluation reports and their results because they do not find these clear aspects.

### *The Demand for Evaluation in Civil Society*<sup>22</sup>

Citizens, social organisations and some political leaders demand evaluation processes with a frequency of 87%, disclosure is done on a smaller scale of 57%. Some of the reasons for this have to do with project control, since social organisations are observers of state actions (República de Colombia, 1991); also because parliamentarians, who follow the development of projects and programmes from the Senate, the House of Representatives and the Municipal Councils, want to know the destination of the items and to what extent they are compliance with, for political purposes, in order to implement other policies or to change the destination of contributions, among the main reasons.

Civil Society's interest in evaluations and their results has been accentuated in the last 20 years. These latest results allow us to infer that the

<sup>22</sup> The reasons written in this section are taken from the conclusions drawn from the questionnaire that was applied in this regard to members of some Social Organisations, some parliamentarians and people who are related to the issues and organisations, whether they are officials of entities, NGO professionals, etc. The survey appears as a file in the appendices of the document.

appropriation and demand for evaluation in Colombia by social organisations has improved. However, as Restrepo points out, “in order to achieve efficient social participation and help democratise the State and society, several principles are proposed: the transparency of all government acts; the incentive of social propositional capacities regarding public policies, without losing its autonomy before the State; positive discrimination in favour of underrepresented social groups; strengthening and construction of social and political actors; institutionalisation of law and participatory practice so that it ceases to be a random and manipulated circumstance; creation of socio-community networks systematising the experiences of participation and; strengthening of population capacities in terms of defining the main variables of political, administrative and economic power” (2001, p. 255). This would guarantee participation and the demand for evaluation and the results of the projects made by the communities and their organisations to a greater extent.

After the signing of the Peace agreement in 2018, this demand is directed, on the one hand, to the attention of human rights, the relocation of the victims of the conflict, the protection of children against abuse, and on the other hand, to the respect of budgets and their uses, destined to support and strengthen community development, the construction of road infrastructure facilitating communication through roads and bridges for small producers in rural areas, Indigenous peoples and communities in a state of extreme poverty. However, “practice shows a totally different situation. The evaluation of social programmes and projects is an infrequent activity, if not exceptional. It is rarely done when projects are prepared and, in general, the social actors involved in them are often reluctant to evaluate what they did” (Cohen & Franco, 1998, p. 82).

### PROFESSIONALISATION SYSTEM OF EVALUATION IN COLOMBIA

The professionalisation of evaluation in Colombia is based on the information emanating from the *Instituto Colombiano para el Desarrollo de la Ciencia y la Tecnología* (Colombian Institute for the Development of Science and Technology, COLCIENCIAS).<sup>23</sup> Its database contains

<sup>23</sup> Constituted by Decree 2869 of 1968 and defined by Decree 585 of 1991, as the Consejo Nacional de Ciencia y Tecnología (National Council of Science and Technology)—COLCIENCIAS. At that time, it assumed the administration of the National

information on the research groups endorsed by this regulatory body, belonging to the Institutions of Higher Education (IES), as well as on certified publications, on the contributions made by the university observatory (which studies the evaluation of public policies in education), on state programmes and projects established in the different development plans and on the *Consejo Nacional de Acreditación* (National Council for Accreditation of university programmes, CNA).

For its part, the *Instituto Colombiano de Fomento de la Educación Superior* (Colombian Institute for the Promotion of Higher Education, ICFES),<sup>24</sup> evaluates education at all levels and investigates the factors that affect its quality.<sup>25</sup> These types of evaluations are made by public or private entities and funding for the development of HEIs derive from their results.<sup>26</sup> The application of State Exams for the Entrance to Higher Education (1964), the SABER tests (1998) and the ECAES (2002) have become, over the years, massive evaluation tests that determine the discourse in the field of quality and evaluation in Colombia and lead to a way of institutionalising evaluation that allows it to be conceived from a political point of view. At the same time, conceptions concerning the educational field such as teaching, learning, knowledge, quality and evaluation itself are standardised, in addition to the consolidation of an evaluation culture in terms of social space, negotiation and transformation, because new relationships between institutions and the State derive from the results, in terms of evaluation; an evaluation policy closely related to initiatives coming from civil society and academia itself is promoted, in addition to the groups that define educational policies.

In this same sense, respecting university autonomy and the freedom of teaching, learning, research and professorship, ICFES supports the refinement of HEIs self-evaluation procedures and, together with the National Council of Higher Education, it proposes to the government mechanisms to evaluate their academic quality. These processes are complemented by external and independent evaluations carried out in educational institutions by academic peers.<sup>27</sup>

System of Science and Technology, today the Ministry of Science, Technology and Innovation (Colciencias, 2016).

<sup>24</sup> Created by Decree 3156 on December 26, 1998 as an independent state-owned company of social nature in the Education sector.

<sup>25</sup> Law 1324 of 2009. Art. 12, Law 1324 of 2009.

<sup>26</sup> Established in Law 635 of 2000, Law 1324 of 2009, Article 12, Law 1324 of 2009.

<sup>27</sup> Coordinated by ICFES, as defined in Law 1324 of 2009, Ent. 0802 Mineducación.



Thus, the State guarantees the quality of the educational service through the practice of inspection and surveillance, for which it conforms the following instances within the Education Ministry (Table 7.2).

**Table 7.2** Instances evaluating service quality in Colombia (own development)

<i>Entity</i>	<i>Goal</i>	<i>Components</i>
Sistema Nacional de Aseguramiento de la Calidad (National Quality Assurance System) (SNIES)	Accountability of HEIs before society and the State Permanent self-evaluation of academic institutions and programmes Culture of evaluation and continuous improvement	Information: SNIES Labour Observatory for Education and Quality Assurance Information System Evaluation: academic peers, CNA, ECAES Promotion: MEN
Comisión Nacional Intersectorial del Aseguramiento de la Calidad de la Educación en Superior (National Intersectorial Commission for Higher Education Quality Assurance) (CONACES)	Applies evaluations to academic programmes for the granting or renewal of Qualified Registration	Evaluation rooms: academic experts
Sistema Nacional de Acreditación (National Accreditation System) (CNA)	It defines policies, strategies, processes and organisations to guarantee that higher education institutions meet the quality standards required for their operation	Public HEIs Private HEIs CNA Academic peers
Consejo Nacional de Acreditación (National Accreditation Council) (CNA)	It fosters high quality conditions in Higher Education institutions. Its academic nature promotes respect for the vision of the IES that undergo the evaluation process to achieve visibility, internationalisation and mobility of their community	Self-evaluation: IES, academic programmes, CNA, External Evaluation: Academic Peers, Final Evaluation: CNA

It is worth noting that there is a difference between institutional evaluation and the accreditation of academic programmes. At the same time, common elements are recognised between both of them, among which it is worth highlighting the role of the academic community

(protagonist/recipient) in the evaluation of institutions and the need for interaction to build knowledge, contribute to the improvement of internal procedures, or to comply with external demands, the latter being those that attract the greatest interest of the actors of the educational community, since institutional accreditation depends on them.<sup>28</sup>

### *Academic Courses of Study and Training*

The report of the Organisation for Economic Cooperation and Development (OECD) explains that one of the greatest achievements of the Educational System in Colombia has been the development of information and evaluation systems that have strengthened evidence-based decision-making (OECD, 2016). In light of this, it is worth noting the country's effort in the implementation of evaluation systems, supported by regulations issued by the State and which permeate not only educational institutions, but have served as a boost to non-profit organisations, foundations and NGOs, whose common objective is to promote evaluation training.

The National Information System for Higher Education (SNIES) registers 118 postgraduate programmes in whose curriculum there are between one and four subjects related to evaluation. To carry out the report, the study plans of the postgraduate courses of 97 accredited public and private universities were reviewed, of which only two master's degrees and 24 specialisations confer the title of 'Master in' or 'Specialist in' evaluation, respectively (Appendix 7.4). At the same time, three HEIs and 16 independent organisations that have been contracted by State entities to carry out the evaluation of public policies, grant the 'Evaluator in' certificate. The latter also develop consultant training programmes, aiming to evaluate academic programmes, financial and educational institutions, accreditation processes or quality management, among others (Appendix 7.5). The data collected on the evaluator training process in Colombia assume that it does not emerge as an intentional plan, within academic programmes. The distribution by disciplines was as follows (Table 7.3)<sup>29</sup>:

<sup>28</sup> See example of the 2019–2020 process at <https://www.mineduacion.gov.co/portal/secciones/Convocatoria-y-concursos/390990> (Process of selection of evaluators for evaluation rooms).

<sup>29</sup> Information taken from <https://hecaa.mineduacion.gov.co/consultaspublicas/ies>.

**Table 7.3** Distribution by discipline (own development)

<i>Discipline</i>	<i>Knowledge area</i>	<i>Percentage (%)</i>
NATURAL SCIENCES	Environment	9.3
AGRICULTURAL SCIENCES	Agriculture	1
SOCIAL SCIENCES	Management	23
	Economics	40
	Education	6
	Law	13
HEALTH AND MEDICAL SCIENCES	Medicine	4.1
ENGINEERING	Industrial Engineering	1

### *Profession/Discipline*

#### *Publications on Evaluation in Colombia*

In Colombia, the indexation of scientific journals depends on the National Bibliographic Index Publindex.<sup>30</sup> Its guidelines are followed by the country's academic and scientific community, to which university publishers and research areas of the different university programmes belong. Nowadays, these publications constitute an evaluation criterion for the accreditation of programmes and respond to the academic exercise of recognised research groups, researchers associated with university institutions. It is a tool for the quality accreditation of academic programmes, teacher selection, salary allocation, management and dissemination of knowledge as one of the CNA criteria.

By 2020, Colombia has 275 A1 indexed journals that are part of the repository of state and private universities. In figures reported by SciELO.org, there are 15,745 published articles (year 2010–2020), related to evaluation in Colombia.<sup>31</sup> Below is the information collected in 14 of the 99 universities with indexed and specialised evaluation journals (14% sample), whose publications are to be found in digital and printed format (Table 7.4). 34% of total articles correspond to the evaluation topic (2010–2020).

<sup>30</sup> See at <https://scienti.minciencias.gov.co/publindex/#/revistasPublindex/historico>.

<sup>31</sup> Distributed in knowledge areas according to OECD provisions: Natural Sciences (7 areas, 48 disciplines)—Engineering and Technology (11 areas, 57 disciplines)—Medical and Health Sciences (5 areas, 60 disciplines)—Agricultural Sciences (5 areas, 14 disciplines)—Social Sciences (9 areas, 29 disciplines)—Humanities (5 areas, 22 disciplines).

**Table 7.4** Universities with indexed and specialised journals in evaluation (14% sample, Universities webpages and Minciencias repository) (own development)

<i>UNIVERSITY</i>	<i>TYPE</i>	<i>#JOURNAL</i>	<i>#EVALUATION</i>	<i>Evaluation articles</i>
Universidad Distrital Francisco José de Caldas <sup>a</sup>	PUBLIC	95	88	491
Universidad Nacional de Colombia <sup>b</sup>	PUBLIC	499	234	947
Universidad del Quindío <sup>c</sup>	PUBLIC	43	111	571
Universidad de Los Andes <sup>d</sup>	PRIVATE	142	76	1642
Universidad del Atlántico <sup>e</sup>	PUBLIC	56	8	17
Universidad de Cartagena <sup>f</sup>	PUBLIC	79	11	3
Universidad de Córdoba <sup>g</sup>	PRIVATE	40	No access	41
Universidad Pedagógica Nacional <sup>h</sup>	PUBLIC	36	54	1276
Universidad Tecnológica de Pereira — UTP <sup>i</sup>	PUBLIC	71	No access	No access
Universidad de La Salle <sup>j</sup>	PRIVATE	36	40	25
Universidad del Magdalena — Uninagdalena <sup>k</sup>	PUBLIC	34	10	10
Universidad de Nariño <sup>l</sup>	PUBLIC	46	20	19
Universidad del Valle <sup>m</sup>	PUBLIC	147	317	145
Universidad Induslilal de Santander— UIS <sup>n</sup>	PUBLIC	88	16	64

(continued)

**Table 7.4** (continued)

<i>UNIVERSITY</i>	<i>TYPE</i>	<i>#JOURNAL</i>	<i>#EVALUATION</i>	<i>Evaluation articles</i>
Universidad Javeriana <sup>o</sup>	PRIVATE	115	27	72

<sup>a</sup> <https://revistas.udistrital.edu.co>

<sup>b</sup> <https://www.hermes.unal.edu.co/>

<sup>c</sup> <https://ojs.uniquindio.edu.co/ojs/index.php/riuq/index>

<sup>d</sup> <https://revistas.uniandes.edu.co/>

<sup>e</sup> <https://www.uniatlantico.edu.co/uatlantico/investigacion>

<sup>f</sup> <https://revistas.unicartagena.edu.co/index.php/index/search>

<sup>g</sup> <https://search.scielo.org>

<sup>h</sup> <https://revistaspedagogica.edu.com>, <http://editorial.pedagogica.edu.co>

<sup>i</sup> no access

<sup>j</sup> <https://ciencia.lasalle.edu.co/>

<sup>k</sup> <http://revistas.unimagdalena.edu.co/>

<sup>l</sup> <https://revistas.udenar.edu.co/>

<sup>m</sup> <http://revistas.univalle.edu.co/>

<sup>n</sup> <https://revistas.uis.edu.co/>

<sup>o</sup> <https://revistas.javeriana.edu.co/>

### *Existence of Professional Organisation VOPE*

Since 1998, Colombia has been part of international evaluation groups (PREVAL). Currently, the country belongs to the International Network for the Evaluation of Public Policies (RIEPP) with 14 members, and it was a founding partner, in 2002, of the Latin American Network for Monitoring, Evaluation and Systematisation of Latin America and the Caribbean (ReLAC), participating in the executive committee (2004–2015). Until 2017, the country coordinated the group Evaluate from Latin America, it was in charge of maintaining the reflection on “the development of evaluation from Latin America and the Caribbean, analysing the relevance and agreement of different approaches, based on our cultural, social and political realities” (ReLAC, 2017).

It also seeks to “(...) generate collective learning and contribute to the construction of Latin American thought on evaluation of public policies, their programmes and projects, from a democratic and contemporary governance perspective” (ReLAC, 2017). Currently, there is no information available to the public about the activities carried out in the last two years.

Other Colombian organisations with international recognition are: *Centro Internacional de Educación y Desarrollo* (International Centre for Education and Development, CINDE), with 40 research projects on evaluation in the field of children and youth; *Centro de Investigación de la Facultad de Educación de la Universidad de los Andes* (Research Centre of the Faculty of Education of the Universidad de los Andes, CIFE) with 17 projects on evaluation and education policy, *Centro de Estudios sobre Desarrollo Económico* (Centre of Studies on Economic Development, CEDE), in the same university<sup>32</sup> with 30 evaluation projects concerning social programmes and projects. Lastly, the *Instituto de Evaluación Tecnológica en Salud* (Institute for Health Technology Evaluation, IETS) with 17 evaluation projects in the health field.

Although an evaluation culture has developed in the last 30 years, in Colombia, the predominant evaluation systems are located in government entities. However, it should be noted that the evaluation practice is carried out bilaterally and with the participation of national and international organisations. Another aspect to highlight is that Colombia does not yet have a body that exclusively certifies evaluators, nor does it have an evaluation training programme other than that offered in the aforementioned university programmes.

### *National Conciliation System*<sup>33</sup>

Our country relies on the following conciliation mechanisms<sup>34</sup>: (a) arbitration, (b) amicable settlement, (c) mediation, finally, (d) conciliation. The General Bureau for Prevention and Conciliation trains the conciliators in turn, endorses other institutions and defines the requirements for the constitution of conciliation centres, legal offices of universities, Chambers of Commerce, which promote leadership in the field of conflict resolution through the houses of justice and conciliation and arbitration

<sup>32</sup> CIFE and CEDE are research and development centres of the Universidad de Los Andes.

<sup>33</sup> The figure of conciliation is conceived as a mechanism of social construction based on dialogue. In other times it has also been known as ‘mediation’, ‘intermediation’, ‘peaceful solution of conflicts’. It is a way of settling differences between people, social groups or states, initially, outside the system or state institutions.

<sup>34</sup> According to Law 446 of 1998.

centres.<sup>35</sup> The current structure of the National System for Conciliation (SNC) and its components is presented in Appendix 7.6.

In Colombia, conciliation has succeeded in incorporating alternative methods of conflict resolution as a common practice of social life. Today, it permeates public and private institutions, communities in rural and urban areas throughout the country, citizens of all strata and conditions, and has become a training space from the earliest grades of pre-school, basic, middle, technical and higher education; it inspires school coexistence manuals and student statutes at different levels of education. The number of cases attended in conciliation rooms between 2003 and 2019, amounts to 80,000 cases per year on average (DANE, 2018).

### *Compliance with Quality Standards and Obligations*

Compliance with the obligations and quality standards in the different public and private organisations is governed by the evaluation principles established by the United Nations Evaluation Group (UNEG). These define evaluation as a systematic and impartial assessment of “an activity, project, programme, strategy, policy, topic, theme, sector, operational area or institutional performance”. Information emanating from the evaluation must be “credible and useful on an empirical basis, facilitating the timely incorporation of the findings, recommendations and lessons into the decision-making processes of organisations and stakeholders”. (UNEG, 2016, p. 10). And ReLAC, in charge of establishing the general principles, procedures, evaluation standards, ethical and behavioural criteria for the practices and professionalisation of the monitoring, evaluation and systematisation of results.

The suggested dimensions and components for evaluation in Latin America and the Caribbean are detailed in Fig. 7.3.

State organisations, universities, foundations, NGOs, consultants, among others, publish the evaluators’ conditions regarding knowledge, education and training, compliance with ethical criteria, application of evaluation competencies and experiences. One of the reference sites in the case of HEIs, is the Curriculum Vitae database of Latin America and the Caribbean, of the National Ministry of CTeI, where academics—as

<sup>35</sup> They emerged in 1995 as multi-agency centres that “bring justice closer to citizens, guiding them about their rights, preventing crime and fighting against impunity” (DNP, 2015).

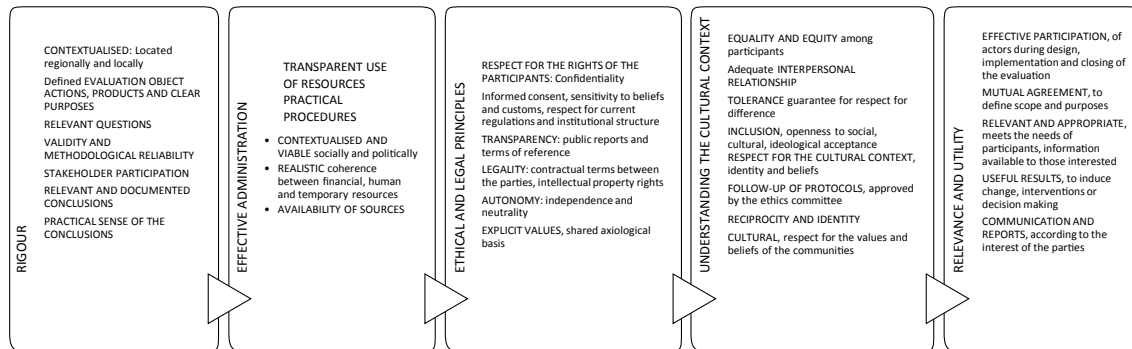


Fig. 7.3 Evaluation dimensions and components (UNEG, 2016)



natural persons—and qualified and accredited research groups appear for evaluation.

For their part, the independent organisations that take part in evaluations consider the indicators established by local, national or international entities to determine, on the one hand, the effectiveness, efficiency and productivity of state policies and, on the other hand, programmes, strategies and evaluation actions that gather the experience of policy implementation. In both cases, monitoring is done by both government entities in the form of Accountability, management reports and periodic bulletins, as well as entities representing civil society linked to the academic sector, observatories, research centres, NGOs, which are joined by international organisations.<sup>36</sup>

Each evaluated and evaluating institution, as well as independent evaluators, commit to the ‘evaluation terms’ covering the different requirements, needs and fields of application. Said terms are mandatory, establishing contractual penalties ranging from pecuniary to legal sanctions (civil and criminal). Just as state entities are held accountable, individuals participating in the evaluation submit reports to boards of directors, departments, or assemblies depending on the type of hiring.

## CONCLUSIONS

In a country that appears in the world as democratic, in addition to laws putting democracy at stake, spaces for citizen participation are important and basic.

The work carried out shows the effort Colombia has made to promote community participation in the formulation and monitoring of public policy, throughout legislation. Until 1991, on the national scene, only Community Action Boards were visible in almost all the municipalities that defended, in some way, the rights of their communities to be present in the multiple development and support plans concerning their well-being and to advocate for their human rights. Despite this, their proposals did not contribute much to improving their quality of life.

Today, there are about 400 social organisations, whose increase is due to the validity of the Constitution of 1991 and in recent years, to the processes of ending the armed conflict specified in the 2018 Peace Agreement. The demand for evaluations and their results, on the part of the

<sup>36</sup> Like IDB, UNICEF, OEI, ECLAC, UNESCO.

social sectors is due, on the one hand, to the defense of human rights, the relocation of the victims of the conflict, the protection of children against abuse, and, on the other hand, to the respect for the budgets destined to support and strengthen community development, and the construction of road infrastructure to facilitate the communication of rural peasants, indigenous peoples and communities living in extreme poverty, through roads and bridges. Although spaces for the discussion of evaluation results and their dissemination through the mass media (television, radio and print) are still missing, it is pertinent to highlight that from the organisation and establishment of MINTIC on, dissemination covers more users in social networks, but connectivity is still very precarious in extreme points of the country.

Finally, it is necessary to recognise that Colombia has grown both regarding the number of organisations making part of evaluation processes and, to a lesser extent, in citizen participation involved. Even so, these processes and their results are strengthened and have given rise to the implementation of actions that demonstrate not only their institutionalisation, but the influence they have, at the moment, as a mechanism for democratic progress and life quality improvement for all citizens represented in the different groups: researchers, academics, members of social action boards, foundations, NGOs, indigenous and agricultural organisations, associations of reintegrated people (former guerrilla members) and educational institutions, among others, whose joint and intentional work with the State consolidate the culture of evaluating programmes and projects, established in the different government development plans throughout the country.

The work started in Colombia to institutionalise evaluation culture must be strengthened in order to: (a) end the distrust of communities to participate; (b) guarantee citizen training in evaluation processes; (c) follow-up and monitor development plans; and d) resignify the common exercise of regulating citizen participation, due to its voluntary nature.

### LIST OF ABBREVIATIONS

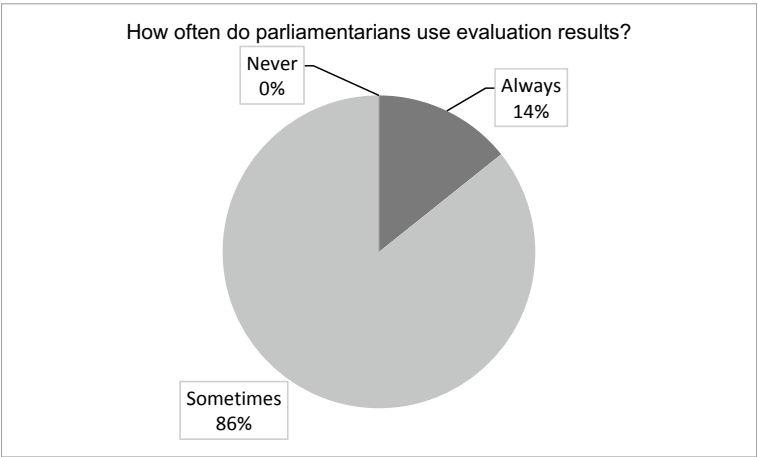
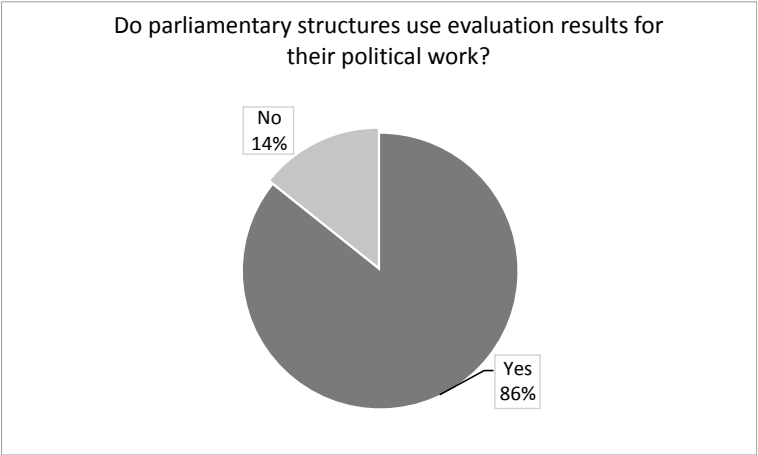
APC	Agencia Presidencial de Cooperación Internacional (Presidential Agency for International Cooperation)
CEDE	<i>Centro de Estudios sobre Desarrollo Económico</i> (Centre for Economic Development Studies)

CINDE	<i>Centro Internacional de Educación y Desarrollo</i> (International Centre for Education and Development)
CIFE	<i>Centro de Investigación de la Facultad de Educación</i> (Research Centre of the Education School)
CTeI	<i>Ciencia, tecnología e Innovación</i> (Science, technology and innovation)
CNA	<i>Consejo Nacional de Acreditación</i> (National Accreditation Council)
COLCIENCIAS	<i>Instituto Colombiano para el Desarrollo de la Ciencia y la Tecnología</i> (Colombian Institute for the Development of Science and Technology)
CONPES	<i>Consejo Nacional de Política Económica</i> (National Council for Economic Policy)
DAFP	<i>Departamento Administrativo de la Función Pública</i> (Administrative Department of the Public Function)
DNP	<i>Departamento Nacional de Planeación</i> (National Planning Department)
DSEPP	<i>Dirección de Seguimiento y Evaluación de Políticas Públicas</i> (Bureau for Monitoring and Evaluation of Public Policies)
GPOR	<i>Gestión por Resultados</i> (Results-based management)
ICFES	<i>Instituto Colombiano de Fomento de la Educación Superior</i> (Colombian Institute for the Promotion of Higher Education)
IDB	Inter-American Development Bank
IETS	<i>Instituto de Evaluación Tecnológica en Salud</i> (Institute of Technological Evaluation of Health)
MIPG	<i>Modelo Integrado de Planeación y Gestión</i> (Integrated Planning and Management Model)
NGO	Non-governmental non-profit organisation
OECD	Organisation for Economic Cooperation and Development
PES	Planeación Estratégica Sectorial (Strategic sector planning)

PREVAL	Programa para el fortalecimiento de la capacidad regional de seguimiento y evaluación para la reducción de la pobreza rural en América Latina y el Caribe (Programme for the strengthening of regional capacities in monitoring and evaluation of rural poverty reduction in Latin America and the Caribbean)
ReLAC	Red Latinoamericana de Seguimiento, Evaluación y Sistematización de Latinoamérica y el Caribe (Latin American Network for Monitoring, Evaluation and Systematisation of Latin America and the Caribbean)
RIEPP	Red Internacional de Evaluación de Políticas Públicas (International Network for the Evaluation of Public Policies)
SGR	Sistema General de Regalías (General Royalty System)
PND	Plan Nacional de Desarrollo (National Development Plan)
SNC	Sistema Nacional para la Conciliación (National System for Conciliation)
SINERGIA	Sistema Nacional de Evaluación de Gestión de Resultados (National Evaluation System of Results)
SMSCE	Sistema de Monitoreo, Seguimiento, Control y Evaluación (Monitoring, Follow-up, Control and Evaluation System)
SIGOB	Sistema de Gestión y Seguimiento a Metas de Gobierno (Management and Monitoring System of Goals of Government)
SIG	Sistema Integrado de la Gestión de Mincomercio (Integrated Management System of Mincomercio)
SNIES	Sistema Nacional de Información de la Educación Superior (National Higher Education Information System)
SISMEG	Sistema de Seguimiento a Metas de Gobierno a corto, mediano y largo plazo a nivel territorial (Monitoring System of Government Goals in the short, medium and long term at the territorial level)
UNEG	United Nations Evaluation Group
WB	World Bank

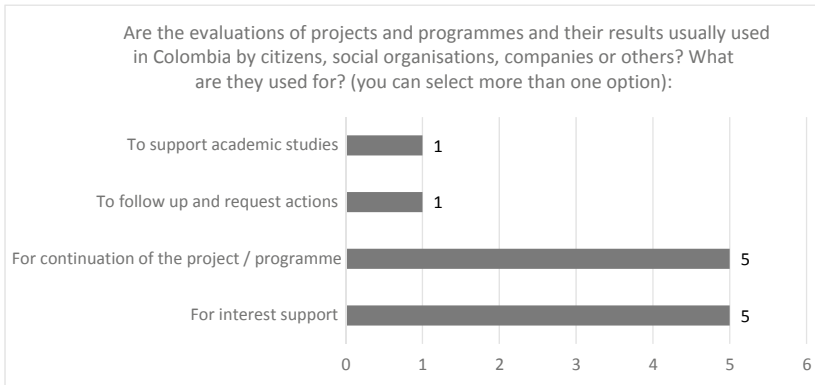
### APPENDIX 7.1

Use and frequency of the evaluation of the parliamentary structure  
(Interview with parliamentarians)



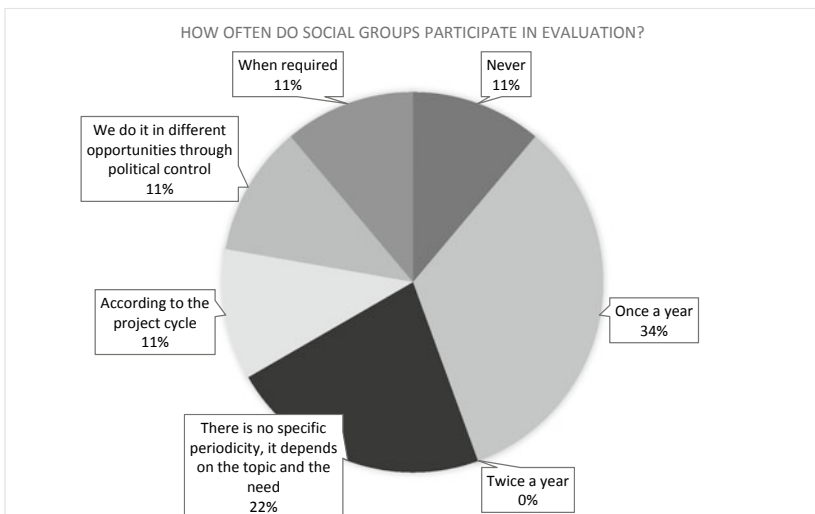
## APPENDIX 7.2

### Purpose of using evaluation results (Interview with officials)



## APPENDIX 7.3

### Frequency of participation in the use of evaluation and its results (Interview with officials)



## APPENDIX 7.4

Programmes that grant a postgraduate degree in evaluation (own development)

<i>Master's degree</i>		
<i>Institution</i>	<i>Programme</i>	<i>Degree</i>
Universidad del Tolima	Master in Environmental Management and Environmental Impact Evaluation	Magíster en Gestión Ambiental y Evaluación del Impacto Ambiental
Universidad Externado De Colombia	Master in Evaluation and Quality Assurance of Education	Magíster en Evaluación y Aseguramiento de la Calidad de la Educación
SPECIALISATIONS		
Universidad del Rosario	Specialisation in Project Evaluation and Development	Especialista en Evaluación y Desarrollo de Proyectos
Universidad Jorge Tadeo Lozano	Specialisation in evaluation of environmental impact of projects	Especialista en evaluación del impacto ambiental de proyectos
Universidad de Antioquia	Specialisation in Socio-Economic Evaluation of Projects	Especialista en Evaluación Socio - Económica de Proyectos
Universidad de Antioquia	Specialisation in Preparation and Evaluation of Private Projects	Especialista en Preparación y Evaluación de Proyectos Privados
Universidad de los Andes	Specialisation in Social Evaluation of Projects	Especialista en Evaluación Social de Proyectos
Universidad de Caldas	Specialisation in Comprehensive Evaluation of Environmental Impacts	Especialista en Evaluación Integral de Impactos Ambientales
Universidad de Caldas	Specialisation in Management and Evaluation of Agricultural and Agroindustrial Projects	Especialización en Administración y Evaluación de Proyectos Agropecuarios y Agroindustriales
Universidad de Medellín	Specialisation in Formulation and Evaluation of Public and Private Projects	Especialista en Formulación y Evaluación de Proyectos Públicos y Privados
Universidad de San Buenaventura	Specialisation in Learning Evaluation	Especialista en Evaluación del Aprendizaje

(continued)

(continued)

<i>Master's degree</i>		
<i>Institution</i>	<i>Programme</i>	<i>Degree</i>
Universidad del Norte	Specialisation in Project Design and Evaluation	Especialista en Diseño y Evaluación de Proyectos
Universidad Industrial de Santander	Specialisation in Project Evaluation and Management	Especialización en Evaluación y Gerencia de Proyectos
Corporación Universitaria del Meta	Specialisation in Project Formulation and Evaluation	Especialista en Formulación y Evaluación de Proyectos
Corporación Universitaria Iberoamericana	Specialisation in Formulation and Evaluation of Social Development Projects	Especialista en Formulación y Evaluación de Proyectos de Desarrollo Social
Fundación Universitaria Autónoma de Las Américas	Specialisation in Project Evaluation and Management	Especialista en Evaluación y Gerencia de Proyectos

## APPENDIX 7.5

Organisations and consultants carrying out evaluations (own development)

<i>No</i>	<i>CONSULTANTS</i>	<i>PROJECTS</i>
1	UNIÓN TEMPORAL ECONOMETRÍA- SISTEMAS DE INFORMACIÓN ESPECIALIZADOS S.A. COLOMBIA SEI	8
2	CENTRO NACIONAL DE CONSULTORÍA	6
3	FEDESARROLLO	2
4	SISTEMAS DE INFORMACIÓN ESPECIALIZADOS S.A. COLOMBIA	11
5	CENTRO NACIONAL DE CONSULTORÍA- EVALUAR	3
6	CONSORCIO FANI CONSULTORES—sin información pública	1
7	GEXPONENCIAL FIRMA CONSULTORA- sin información pública	1
8	ECONOMETRÍA CONSULTORES	97
9	GEXPONENCIAL—CENTRO INTERNACIONAL DE EDUCACIÓN Y DESARROLLO HUMANO CINDE	1
10	PROYECTAMOS COLOMBIA SAS (consultants)	1

(continued)

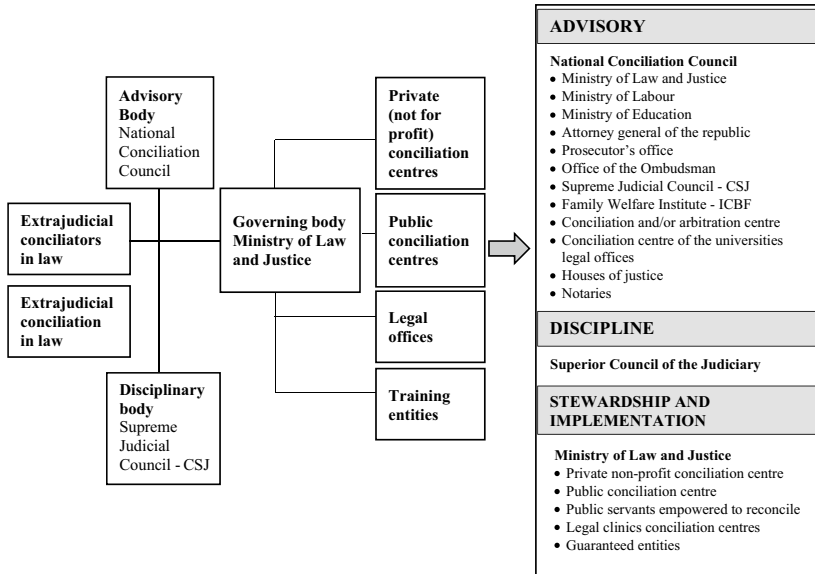


(continued)

<i>No</i>	<i>CONSULTANTS</i>	<i>PROJECTS</i>
11	CONTRATISTA ERNST & YOUNG S.A.S—Colombia	1
12	FACULTAD DE ECONOMÍA, UNIVERSIDAD DE LOS ANDES CENTRO DE ESTUDIOS DE DESARROLLO ECONÓMICO-CEDE	1
13	INNOVOS GROUP	1
14	INSTITUTO DE INVESTIGACIÓN Y DESARROLLO EN ABASTECIMIENTO DE AGUA, SANEAMIENTO AMBIENTAL Y CONSERVACIÓN DEL RECURSO HÍDRICO, (CINARA) Universidad del Valle	1
15	CENTRO DE ESTUDIOS DE DESARROLLO ECONÓMICO (CEDE) School of Economics, Universidad de Los Andes	30
16	CIFE, Universidad de Los Andes	17
17	PROES S.A INGENIEROS CONSULTORES	1
18	E—VALUA	13
19	CINDE	40
20	IETS	16

## APPENDIX 7.6

Structure of the National Conciliation System (DNP, 2016, pp. 113–115)



## REFERENCES

- Circular 62. (2013, September). [https://xperta.legis.co/visor/temp\\_legcol\\_1ea\\_c3986-b1b4-414f-8977-b816baed919e](https://xperta.legis.co/visor/temp_legcol_1ea_c3986-b1b4-414f-8977-b816baed919e). Accessed on 13 April 2020.
- Cohen, E., & Franco, R. (1998). *Evaluación de proyectos sociales*. Siglo XXI.
- Colciencias. (2016). *Guía para el reconocimiento y medición de grupos de investigación e investigadores*. Departamento Administrativo de Ciencia, Tecnología e Innovación—Colciencias Dirección de Fomento a la Investigación. [https://minciencias.gov.co/sites/default/files/ckeditor\\_files/guia-reconocimiento-y-medicion-de-grupos-e-Investigadores.pdf](https://minciencias.gov.co/sites/default/files/ckeditor_files/guia-reconocimiento-y-medicion-de-grupos-e-Investigadores.pdf). Accessed on 1 March 2020.
- CONPES 2688. (1994, January). *Evaluation of public sector results at national level*. <https://colaboracion.dnp.gov.co/CDT/Conpes/Economicos/2688.pdf>. Accessed on 31 March 2020.

- CONPES 2790. (1995). *Results-oriented public management*. <https://www.armada.mil.co/es/content/documento-conpes-2790-de-1995-gestión-pública-orientada-resultados>. Accessed on 31 March 2020.
- CONPES 3294. (2004, June). *Renewal of public administration, results-based management and reform of the National System for Management Evaluation and Results (SINERGIA)*. <https://colaboracion.dnp.gov.co/CDT/Conpes/Económicos/3294.pdf>. Accessed on 15 March 2020.
- CONPES 3515. (2008, April). *Strengthening of public information, Monitoring and evaluation for results-based management in Colombia*. <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/3515.pdf>. Accessed on 13 April 2020.
- DANE. (2018). *Resultados Censo Nacional de Población y Vivienda*. <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/censo-nacional-de-poblacion-y-vivienda-2018/cuantos-somos>. Accessed on 15 June 2020.
- Decree 3156. (1968, December). [https://www.mineduacion.gov.co/1759/w3-article-104221.html?\\_noredirect=1](https://www.mineduacion.gov.co/1759/w3-article-104221.html?_noredirect=1). Accessed on 31 March 2020.
- Decree 3152. (1989, March). <http://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Decretos/30036303>. Accessed on 15 March 2020.
- Decree 2167. (1992, December). [https://www.redjurista.com/Documents/decreto\\_2167\\_de\\_1992\\_dnp\\_-departamento\\_nacional\\_de\\_planeacion.aspx/](https://www.redjurista.com/Documents/decreto_2167_de_1992_dnp_-departamento_nacional_de_planeacion.aspx/). Accessed on 15 March 2020.
- Decree 695. (2003, March). <http://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Decretos/1148647>. Accessed on 1 June 2020.
- Decree 3517. (2009, September). <http://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Decretos/1529717>. Accessed on 12 April 2020.
- Decree 1832. (2012, August). <http://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Decretos/1358237>. Accessed on 15 March 2020.
- Decree 414. (2013, March). <http://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Decretos/1851095>. Accessed on 12 April 2020.
- Decree 1163. (2013, March). <http://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Decretos/1235332>. Accessed on 13 April 2020.
- Decree 817. (2014, April). <http://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Decretos/1171745>. Accessed on 31 March 2020.
- Decree 1083. (2015, May). <http://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Decretos/30019920>. Accessed on 16 April 2020.
- Decree 2189. (2017, May). <http://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Decretos/30019920>. Accessed on 16 March 2020.
- Departamento Administrativo de la Función Pública (DAFP). (2017). *Modelo integrado de planeación y gestión*.
- Departamento Nacional de Planeación (DNP). (2010). *Evolución de SINERGIA y evaluaciones en administración del Estado. 15 Años del Sistema Nacional*

- de Evaluaciones de Gestión y Resultados–SINERGIA*. Una mirada desde las evaluaciones de política pública más relevantes. Bogotá D.C., Colombia.
- Departamento Nacional de Planeación (DNP). (2015). *Análisis Conceptual del Sistema Nacional de Conciliación en Colombia en sus 25 años: Construyendo diálogo y paz para el futuro*. Bogotá D.C., Colombia. <https://bibliotecadigital.ccb.org.co/bitstream/handle/11520/14095/25%20a%C3%B1os%20del%20Sistema%20Nacional%20de%20Conciliaci%C3%B3n.pdf?sequence=1>. Accessed on 09 May 2020.
- Departamento Nacional de Planeación (DNP). (2016). *Análisis Conceptual del Sistema Nacional de Conciliación en Colombia en sus 25 años: Construyendo diálogo y paz para el futuro*. Bogotá D.C., Colombia. <https://bibliotecadigital.ccb.org.co/handle/11520/14095>. Accessed on 16 March 2020.
- Departamento Nacional de Planeación (DNP). (2019). *La institucionalización de la evaluación en Colombia–SINERGIA*. Bogotá D.C., Colombia.
- Hernández, R., Hernández, Y. F., Gil, M., & Cárdenas, B. E. (2018). Evaluación del modelo integrado de planeación y gestión (MIPG) en las entidades territoriales del estado colombiano. *Revista Aglala*, 9(1), 443–463.
- Law No. 134. (1994, May). <https://pdba.georgetown.edu/Electoral/Columbia/ley134-94.html>. Accessed on 1 June 2020.
- Law No. 489. (1998, December). <http://www.suin-juriscol.gov.co/viewDocumento.asp?ruta=Leyes/1832980>. Accessed on 16 March 2020.
- Law No. 635. (2000, December). <http://www.suin-juriscol.gov.co/viewDocumento.asp?id=1689170>. Accessed on 17 May 2020.
- Law No. 850. (2003, November). <http://www.suin-juriscol.gov.co/viewDocumento.asp?ruta=Decretos/1148647>. Accessed on 1 June 2020.
- Law No. 1257. (2008, December). [https://www.oas.org/dil/esp/LEY\\_1257\\_DE\\_2008\\_Colombia.pdf](https://www.oas.org/dil/esp/LEY_1257_DE_2008_Colombia.pdf). Accessed on 1 June 2020.
- Law No. 1324. (2009, July). <http://www.suin-juriscol.gov.co/viewDocumento.asp?ruta=Leyes/1677369>. Accessed on 15 March 2020.
- Law No. 1530. (2012, May). <http://www.suin-juriscol.gov.co/viewDocumento.asp?ruta=Leyes/1682780>. Accessed on 16 March 2020.
- Law No. 1753. (2015, June). <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=61933>. Accessed on 16 April 2020.
- Legislative Act No. 05. (2011, July). <http://wsp.presidencia.gov.co/Normativa/actos-legislativos/Documents/2011/ActoLegislativo-05-18julio2011.pdf>. Accessed on 16 March 2020.
- OECD. (2016). *Education in Colombia–Highlights 2016*. <http://www.oecd.org/education/school/Education-in-Colombia-Highlights.pdf>. Accessed on 24 April 2020.
- Pérez Yarahuán, G., & Maldonado Trujillo, C. (Eds.). (2015). *Panorama de los sistemas nacionales de monitoreo y evaluación en América Latina*. México.

- ReLAC. (2017). *Grupos de trabajo*. ReLAC. <https://www.relac.net/grupos-de-trabajo/>. Accessed on 10 March 2020.
- República de Colombia. (1991). *Constitución Política de Colombia*. Chapter II. De los planes De desarrollo (pp. 341–344).. <http://wsp.presidencia.gov.co/Normativa/Documents/Constitucion-Politica-Colombia.pdf>. Accessed on 10 March 2020.
- Resolution 63. (1994). [https://colaboracion.dnp.gov.co/CDT/Sinergia/Documentos/DNP\\_Resolucion63\\_1994.pdf](https://colaboracion.dnp.gov.co/CDT/Sinergia/Documentos/DNP_Resolucion63_1994.pdf). Accessed on 16 April 2020.
- Restrepo, D. I. (2001). Participación Social: Relaciones Estado-Sociedad Civil. *Revista De Salud Pública*, 3(3), 245–267.
- Roth Deubel, A. N. (2009). La evaluación de políticas públicas en Colombia: Una mirada crítica a partir de las prácticas evaluativas oficiales de los programas de la " Red de Apoyo Social". *Revista Del Clad Reforma y Democracia*, 45, 161–186.
- Serra, A., Figueroa, V., & Saz, Á. (2007). Modelo abierto de gestión para resultados en el sector público. *Banco Interamericano de Desarrollo (BID)*. Centro Latinoamericano de Administración para el Desarrollo (CLAD). Santiago.
- Statutory law No. 1757. (2015, July). [http://www.secretariasenado.gov.co/senado/basedoc/ley\\_1757\\_2015.html](http://www.secretariasenado.gov.co/senado/basedoc/ley_1757_2015.html). Accessed on 1 June 2020.
- UNEG. (2016). *Norms and standards for evaluation (Spanish)*. [https://www.unicef.org/evaluation/files/UNEG\\_NormsandStandards\\_for\\_Evaluation\\_Spanish\\_2017\\_\(1\).pdf](https://www.unicef.org/evaluation/files/UNEG_NormsandStandards_for_Evaluation_Spanish_2017_(1).pdf)Accessed on 29 May, 2020.



## Evaluation in Costa Rica

*Mayela Cubillo Mora and María Camila León Betancourth*

### INTRODUCTION

The institutionalisation process of the evaluation culture in Costa Rica has been underway for several decades, and there have been numerous advances and achievements, but also some setbacks. Up to the present day, it could be affirmed that the country continues in the middle of the process of institutionalising practices and strategies that promote the evaluation culture in state institutions and in Costa Rican society in general, this is a reason to expect several challenges and tasks to perform in order to consider that the institutionalisation of evaluation is completed.

To see this with greater clarity, it is necessary to make a brief historical route by the different processes that the country has gone through before the present practices being developed to promote the institutionalisation of evaluation. After the above-mentioned historical process brief review, each section of the chapter will allow to illustrate the particular

---

M. Cubillo Mora (✉)

Faculty of Economics, University of Costa Rica, San José, Costa Rica

e-mail: [mayela.cubillo@ucr.ac.cr](mailto:mayela.cubillo@ucr.ac.cr)

M. C. León Betancourth

Universidad Nacional de Colombia, Bogota, Colombia

facts, challenges and opportunities that Costa Rica has in the matter of institutional processes, diffusion and entailment of the civil society, and professionalisation of evaluation.

The long history of evaluation in Costa Rica began in the mid-twentieth century in a problematic context of public resource management and administration in which the Ministry of Finance was the only figure in charge of formulating the national budget (Fonseca, 2001). This previous context was changed by means of the Political Constitution (1949) which created the Comptroller General's Office of the Republic (Contraloría General de la República, CGR) as an autonomous body attached to the legislative assembly, and with the role to exert control of legality in public management. The creation of the CGR is the first sign of interest on the part of the Costa Rican State to count on control mechanisms in place, given that this entity began with the process of control and evaluation of resources in the country. In March 1993, the Department of Protection of Citizens' Rights of the Republic was also created as an auxiliary body of the legislative assembly with the responsibility for human rights. The processes of evaluation of public policies, plans, programmes and projects are regulated in what is explained later in this document.

Subsequently, in the year 1963, the planning office was created, which would become the Ministry of National Planning and Economic Policy (Ministerio de Planificación y Política Económica, MIDEPLAN) in May 1974, as an organisation attached to the presidency of the republic, and which emerged because of the development model of service provision that was available at that time (Fonseca, 2001). A year later, the law of National Planning (Law N°5525 1974) would be responsible for shaping the national planning system, and the MIDEPLAN as its rector, giving the latter responsibility for systematically and continuously evaluating the results obtained from the implementation of policies, programmes and projects.

Even though in Costa Rica international organisations expressed the need to have an entity in charge of evaluating public interventions, it was not until 1994 that this materialised. Thanks to a rethinking process of the economic model and the state structure that sought to modernise the Costa Rican state (García & Ugalde, 2015), spaces that fomented the political, legal and administrative control were created, they were illustrated by the model of public management by results, in which the creation of the National Evaluation System (Sistema Nacional de Evaluación, SINE) occurs, as a management mechanism to be used by the

presidency of the republic to evaluate the actions of the public sector and to feedback the decision-making process. The SINE is formed by the presidency of the republic, MIDEPLAN (who coordinates it), Ministry of Finance, Institutional Planning Units and the sectoral governing bodies at the executive power level (MIDEPLAN, 2019).

The Evaluation and Monitoring Area (Asociación Centroamericana de Evaluación, AES) was created as the body in charge of managing and coordinating SINE. Among the functions of the AES are the monitoring and evaluation of compliance of goals of the National Development Plan (Plan Nacional de Desarrollo, PND), the creation of strategies and methodological guidelines on evaluation, and the promotion of a culture of evaluation in the country.

The conceptual and methodological design of the SINE occurred between 1994 and 1995 and counted on the technical support of the office of evaluation of the Inter-American Development Bank, and on compared cases of countries like Colombia, Chile, Canada, Australia and the United States (García & Ugalde, 2015). It is from this perspective that a new culture, based on evaluation for accountability, was applied in Costa Rica.

In general, the SINE was conceived as a tool to strengthen government management, and to promote the strengthening of the management capacity of the public sector by means of three elements that include measurement, the promotion of quality results, the generation of information for more efficient use of public resources and the availability of information on monitoring and evaluation in different state interventions (García & Ugalde, 2015).

The operation of the SINE was then structured under two modules: Self-assessment and strategic evaluation. The ministries and executive presidents are responsible for carrying out the process of self-evaluation that has the annual operating plan, the annual work plan and the budget of the institutions as instruments. This process aims to develop capacities in evaluation at different institutional levels, in addition to providing spaces for monitoring and accounts settlement on the priorities contained in the PND of the current government, or any other plan, policy or project of the State in general.

On the other hand, strategic evaluation counts on MIDEPLAN as the coordinating and facilitating entity and has the commitments by results that evaluate policies, plans, programmes and projects of public interest and that are determinant for the goals and objectives contained in the



PND as instruments. In addition to the above, the SINE has the operation of institutional planning units that serve as sectoral and institutional links and are in charge of carrying out evaluations within the institutions.

Both AES and SINE focused their efforts on the monitoring and achievement of goals by the PND for more than 16 years. Among other things, the approval of the law on Financial Administration of the republic and Public Budgets (Law N°8131, 2001) allowed a change in the vision of the SINE and brought it closer to the evaluation approach with which it currently works, by linking institutional planning with budget formulation.

In addition, the law on Financial Administration introduced new dynamics for monitoring and evaluation processes in the country, increasing the coverage of the SINE to more than 100 institutions, and creating a link between the PND and the Institutional Operational Plans (Plan Operativo Institucional, POI) strengthening the planning, budgeting, monitoring and evaluation instruments of the institutions.

Law N°8131 indicates in its article 26 that ‘The Financial Administration System of the public sector will be made up of the set of norms, principles and procedures used, as well as by the entities and bodies participating in the planning, obtaining, allocation, utilisation process, registration, control and evaluation of their financial resources’, which ensures the PND-POI-Budget relationship.

Starting in 2010, MIDEPLAN would fully enter into the processes of strategic evaluation of policies, plans and projects in the absence of a consolidated practice in the country, which would be reflected in diverse efforts materialised in the formulation of guidelines and methodologies to guide the evaluation processes. One of those efforts resulted in the creation of the annual matrix of sectoral and institutional programming, follow-up and evaluation, as a technical instrument that determines guidelines and methodologies for the planning and monitoring of the PND goals.

This was followed in 2012 by the promulgation of the MIDEPLAN (2012), which also made it possible to define the theoretical and conceptual approach, as well as the methodological framework to guide the process of institutionalisation in the country. The manual aims to improve public management by means of sustained decision-making through a strategic assessment of the design, implementation and results of policies, programmes and projects impelled by the SINE.

The consolidation process of MIDEPLAN as the lead governing agency on the subject of evaluation in the country occurred hand in hand with the Project for Capacity Building in Evaluation (FOCEVAL) in its first stage as of 2011, through which they began to generate spaces for training, financing and technical assistance for the implementation of evaluations through the ‘learning by doing’ methodology. It is thanks to this that for the first time in the country the PND 2011–2014, MIDEPLAN (2010) includes the development of four evaluations under the attention of AES of MIDEPLAN as one of its goals.

The attainment of these four evaluations, besides providing the MIDEPLAN with experience on the subject of evaluation, resulted in a recommendation on the formulation of a National Agenda of Evaluations to consolidate the approach of management by results in the country’s public sector. This is why in 2014 the first National Evaluation Agenda (Agenda Nacional de Evaluación, ANE) was created as a milestone within the SINE, since it implies a substantial advance towards systematisation in the evaluation processes (García & Ugalde, 2015).

ANE has 15 public interventions distributed in nine sectors, which are also contained in the PND 2015–2018, MIDEPLAN (2014) by means of a structuring according to the organisation of the executive authority; in addition to involving interventions identified as strategic by different sectors and institutions from the public sector. The agenda also directly involves the institutions or organisations that oversee selected projects, given that they themselves carry out the evaluations, accompanied and advised by the MIDEPLAN’s AES, responsible for providing methodologies, strategies and technical guidelines for their development. At present, some of these projects have successfully concluded their phase of evaluation and dissemination of results, whereas in other cases they were expected to conclude in the last months, along with the expiration of the PND.

At the moment, Costa Rica counts on the MIDEPLAN as a coordinating and governing entity of the planning and evaluation processes, materialised by means of the SINE. In addition, it has fortified the bonds that link the MIDEPLAN with the Ministry of Finance, for the emission of monitoring guidelines to the PND, strategic evaluation and budgetary programming. In addition to this work, the work of the general comptroller of the republic, as the entity that oversees public finances, also monitors compliance with and progress in the fulfilment of the objectives contained in the PND. The work of these three institutions allows for the

preparation and joint assessment of reports that are submitted to different instances, such as the legislative assembly, and in which the actions taken by the government are measured and evaluated.

This historical overview of the country's progress in the evaluation of policies, programmes and projects shows that the main and greatest efforts made by governments and public institutions in the area of evaluation, have focused on monitoring and evaluating the objectives contained in the current PND. As it will be seen below, in recent years, the MIDEPLAN has continued its work of follow-up and monitoring of the PND, but also has provided with opportunities for the development of evaluations in other subjects, which has allowed evaluation to reach different agencies and is currently in a more extensive stage.

Finally, it will be seen that one of the greater challenges that the country faces, especially at the level of institutional hierarchies, is to change from a culture of accountability to one based on evaluation as a tool for improving public management.

## INSTITUTIONAL STRUCTURES AND PROCESSES (POLITICAL SYSTEM)

### *Evaluation Regulations*

Costa Rica is governed by the political constitution of November 7, 1949, which establishes a presidential system and a unitary state, with a division of three branches: legislative, executive and judicial. MIDEPLAN is part of the executive branch, which holds the stewardship of the national planning policy and its evaluation, so that its scope is the entire political system.

To date, Costa Rica is formulating its national policy on evaluation, which will make it possible to direct the efforts of the government and the institutions towards the institutionalisation of these processes; the initiative is being presented to different sectors through workshops and working groups and is expected to be presented at the national level by the end of 2018.

Despite this, the country has a long legal history that began in the 1970s which has formed the basis of what is now the SINE under MIDEPLAN. The legal basis for evaluation in Costa Rica lays its foundation on the 1949 constitution, which was amended in article 11 by Law N°8003

(2000) in order to incorporate evaluation by results and accountability as inspiring principles of public management.

“In the broad sense of the term, public administration shall be subject to a procedure of performance evaluation and accountability, with the consequent personal responsibility for the civil employees in the fulfilment of their duties. The law shall establish the means for this control of results and accountability to operate as a system that covers all public institutions”, (Political Constitution, 1949, art. 11).

In addition, the national planning law (1974) grants MIDEPLAN, the power to monitor and evaluate the progress and objectives of the PND; in its article two, the law imposes on MIDEPLAN the role and responsibility of systematically and permanently evaluating the results of government plans, programmes, policies and projects. As mentioned above, MIDEPLAN and the national planning law were conceived in principle to follow up almost exclusively on the goals of the PND, so its role was focused on this until almost the early 2000s, when the Ministry broadened its vision and moved towards an ampler sense of evaluation.

For this reason, in 1994, MIDEPLAN created the AES through the executive decree N°23,323 (1994). The AES emerges as part of the Ministry’s organisational structure and has the responsibility of monitoring and evaluating the goals of the PND, which is why it is constituted as the coordinating entity of the evaluation and monitoring processes in the country. That same year the SINE was created by executive decree N°23,720 (later amended by executive decree N°35,755 of 2010) as a tool to observe government actions and promote a strategic use of public resources with a view to articulating efforts around the priorities of the PND.

Seven years later, the approval of the law on Financial Administration of the Republic and Public Budgets (2001), established in article four that public budgets must respond to the institutional operational plans, and that these in turn must aim to comply with the PND.

The importance of this law lies in its capacity to bind to the processes of budgeting, planning and evaluation of goals of the national plan, as well as in its ability to link the actions of three central institutions for evaluation in the country, such as MIDEPLAN, the Ministry of Finance and the CGR. For this reason, articles 52, 55 and 56 of the law stipulate that the treasury and the MIDEPLAN must evaluate and prepare reports on the fulfilment of goals and the behaviour of public finances, and that

these reports are presented to the CGR as the entity in charge of oversight in the country.

The law of Financial Administration also has regulations (executive decree N°32,988, 2006) specifying the types, timing and recipients of the reports to be drafted. It would be expected that this administrative regulation would be a key instrument to define roles and tasks in the matter of evaluation in the country, or at least for the three institutions directly involved, but in practice there are gaps in the role that each institution should play in the process, which has led to duplication of information within the process; it is also noted that although the law imposes evaluation responsibilities on the Ministry of Finance and the CGR, these institutions do not count on the technical and professional mechanisms to comply with the methodological requirements of an outcome evaluation, which has directed the processes towards mechanisms of accountability and achievement of goals.

Finally, in 2011 the executive decree N°36,901 (2011) was issued for the creation of the Inter-institutional Technical Coordination Commission for Planning, Programming and Evaluation of the Institutional and Sectoral Management of the Executive Authority. The commission was raised as an entity to narrow bonds between the MIDEPLAN and the Ministry of Finance as institutions responsible for planning and budgeting, respectively. It also sought to define guidelines, instruments and concepts that would allow for greater monitoring to the compliance of the PND.

Later, in 2014, and with the intervention of the FOCEVAL project, the figure of the commission was modified to what is now known as the national evaluation platform, where other organisations besides the MIDEPLAN and the Ministry of Finance, such as the Department of Protection of Citizens' Rights and the Legislative Assembly participate, as well as representatives of civil society, evaluation networks and the academy by means of the participation of the Centre for Research and Training in Public Administration (Centro de Investigación y Capacitación en Administración Pública, CICAP) and the Master's programme in evaluation of development programmes and projects of the University of Costa Rica (UCR). Among the main roles of the national platform, one is to support initiatives and the other to give sustainability to the development of evaluations in the country; in addition, it is constituted as an open space where all public or private organisations, interested in the subject, can participate. The platform has become a consolidated space

of communication and coordination between different actors who make particular efforts in the subject of evaluation, and has served as a space to contribute to the formulation of the national policy of evaluation that is being constructed. After this, in 2016, the regional evaluation platform was created.

Likewise, through the interviews conducted, it was observed that institutions such as the CGR and the Ministry of Finance have a clear understanding of their roles in evaluation, but they do not count on the technical and professional apparatus to carry out procedures that fulfil all the technical and methodological requirements that an evaluation demand. While MIDEPLAN, as the lead agency for the process, and having received strong support from the FOCEVAL project in terms of training and education in the recent years, is now more prepared to face evaluation processes technically speaking. Therefore, one of the main challenges that the country has is the precise definition of the roles and the scope of the evaluations within the public sector, which would be expected to be clarified by the national evaluation policy. It also highlights the need to strengthen the national evaluation capacities beyond the MIDEPLAN, with specific emphasis on institutions that by law must carry out such processes, like the Ministry of Finance.

On the other hand, some progress has been made by the legislative assembly to include evaluations and the use of results by parliamentary structures. To this end, the assembly and MIDEPLAN have worked together to build bridges and build awareness of the project through training and awareness-raising processes carried out by the FOCEVAL Project.

Even though at the moment, the legislative evaluation is not an institutionalised practice, since it does not count on a formal agreement that endorses it, the possibility and capacity for the accomplishment of ex-ante, mid-term and ex-post evaluations of the country law have been explored (E. Paniagua, personal communication, June 14th, 2018). Currently, such processes have the direct participation of the library of the legislative assembly, the Centre for Legislative Research and the Centre for Statistical Information, as well as the participation of the services sector.

Within this sector, there is the possibility of working on ex-ante evaluations of pre-legislative studies carried out by the library, which are aimed at generating an argumentative research to accompany the proposals of the deputies, before these become a concrete project of law. At the moment there is continued and deep analysis of this type of evaluations

since their methodology and recommendations do not adjust technically to the parameters of an evaluation. On the other hand, the area of technical services has explored the possibility of making mid-term evaluations of programmes in progress through a consolidated role based on the accomplishment of technical reports to the projects of law presented by the assembly deputies.

Finally, ex-post evaluation has been carried out with the help of the Chilean and German experience, which has allowed the procedure to be enriched and to generate the first pilot plan for the ex-post evaluation of a law in Costa Rica, which is expected to be completed by the end of July 2017. The implementation of the pilot plan on ex-post evaluation has also benefited from the support and advice of the FOCEVAL project and the MIDEPLAN and is expected to serve as a mechanism to influence decision-making within the assembly.

Therefore, it can be affirmed that by the accomplishment of evaluations and the use of results, a process of incorporation of evaluations within the legislative assembly as well as within the parliamentary structures is initiating.

At present, and thanks to the interviews conducted, it cannot be concluded that the evaluations are used precisely by members of parliament for decision-making purposes, given that the pre-legislative studies that would constitute the ex-ante evaluation of laws are voluntary as is the inclusion of the findings of the technical reports that constitute the mid-term evaluation. It is also necessary to mention that the information and awareness processes within the assembly, although they have taken place, are still vulnerable to political processes, therefore, and given the recent change in the assembly, it would be expected that the information and training efforts on evaluation will begin to take place from next year on.

### *Evaluation Practices*

Costa Rica has SINE as a national entity attached to MIDEPLAN, which is responsible for evaluating government management to promote management capacity in the public sector. The work of the SINE is accompanied by institutional planning units, as organisms within each institution or ministry, which are developed as sectoral links in charge of carrying out evaluations. Despite the fact that the roles of the institutional planning units are well defined, it was found that in practice, most of them

do not have trained personnel in evaluation. Therefore, the processes of follow-up, monitoring and evaluation of the institutional operational plans and strategic plans of the institutions, which oversee these units, have been carried out in an empirical manner. It is not surprising then that most of these procedures in public institutions do not meet either the methodological or the technical requirements specified by an evaluation of any type. The lack of knowledge and experience on evaluation has therefore led to a much more focused approach to accountability and internal auditing, highlighting specific institutions' expenditures and specific objectives.

The evaluations carried out in the framework of the ANE have been financed by the public budget, in general, they have been evaluations contracted to external suppliers and in other occasions they have been learning-doing processes with external evaluators and MIDEPLAN personnel. Some cooperation agencies such as German Corporation for International Cooperation GmbH (GIZ) have carried out evaluations on their own, such as those related to the competitiveness and environment programme.

Accompanying this, a worrying level of lack of knowledge and awareness among decision makers was observed. In general, the interviews carried out allow us to conclude that evaluation does not yet have a positive perception in this area, given that it tends to be thought that the nature of the evaluation refers to the measurement of skills for a position, or the development of personal roles that could end in a sanction, or even retirement of the occupied position. Therefore, the challenge in this aspect continues to be the change of the perspective of evaluation into a positive one, so it can be understood as a tool to improve the public administration, mainly on behalf of decision makers and managers who are in charge of incorporating evaluation results and recommendations. Evaluations of process and results (outcomes and impacts) and even the development of baselines are still difficult for the interviewees to observe, which can be attributed to the lack of knowledge regarding their components and benefits, which is why visualise the evaluation as audit processes and accountability.

As for the scope of completed evaluations, the exercise of the national evaluation agenda could be considered as the greatest evaluation effort in the country, which also had the technical and methodological necessary advice to comply with the standards of evaluation.



The first national evaluation agenda, constituted in PND 2015–2018 (2014), formulated the accomplishment of 15 evaluations of projects and national programmes of different sectors under the responsibility of diverse public institutions. Similarly, it has been the greatest exercise to equip MIDEPLAN with the needed experience as well as to other participant institutions on the formulation of terms of reference and recruitment of evaluators. As this is the first national evaluation exercise, ANE does not yet have an established frequency, but it is expected to be carried out with each PND.

Among the challenges facing a forthcoming national agenda is its expansion into different sectors, which are interested in carrying out evaluations and may represent benefit for the government, sectoral, municipal projects or even projects that are proposed by civil society, not necessarily public administration, are some of the areas in which the agenda could intervene. Likewise, it is necessary to effectively get civil society involved, since they did not have any participation in the formulation of the first agenda, it would be hoped that this would change according to the recent linkage of civil society to the national platform and the construction of the national evaluation policy. Finally, another challenge for the agenda is that its formulation and development is not depending on the political will of each current government, given that is not yet a consolidated practice, it is hoped that the agenda will continue to be developed with each incoming government, so that it can be established as an institutionalised practice in the country.

### *Use of Evaluations*

As it was previously mentioned, some sectors have had the opportunity to work in a more direct way with the MIDEPLAN and have counted on the advising of the FOCEVAL Project; but other sectors, both within and outside the public administration, have not had continuous training processes on this matter. This is why it is not possible to distinguish an overall picture in terms of evaluation skills and use of results, given that the level of knowledge and development of practices varies considerably from one sector to another. Often, the lack of evaluation or the poor use of its results is not due to a lack of interest or commitment, but due to the lack of knowledge about evaluation and of financial resources to hire evaluators.

Among the closest sectors to MIDEPLAN that have been interviewed, which therefore have a higher level of training and awareness, are the legislative assembly, the Ministry of Agriculture, the Department of Protection of Citizens' Rights, and the office of the comptroller general of the republic; this does not mean that the follow-up monitoring an evaluation processes carried out with the above-mentioned institutions comply with the technical and methodological requirements of an evaluation, given that, as mentioned above, many of the efforts are guided to comply with accountability procedures.

Other sectors interviewed have not had training processes on evaluation, have not had the opportunity to conduct an evaluation within their institutions, or do not have fully established formal mechanisms for conducting evaluations. These sectors are the Costa Rican tourism board, the Ministry of Health, and civil society organisations at the grassroots or community level.

On the other hand, there is a low level of use of evaluation results especially when considering that the impact of the results and its involvement in concrete actions depend on the decision of the heads of the institutions, and therefore on political will, since its use is not mandatory.

In scenarios such as civil society, the conduct of evaluations and the use of results are considerably reduced, especially given the low level of knowledge, education and training in this regard. The small sector in civil society organisations that is conducting evaluations, generate alternative or shadow reports on the country's situation on a particular issue; the resources for the construction of such reports depend heavily on international cooperation, or almost always on foreign donors, and are very focused on accountability and goals achievement.

On the other hand, a mechanism to guarantee the use of the results of the evaluations, at least in public administration institutions, is through the involvement of these results and their action plans in the annual operational plans of each institution; besides to assure that the results are tied to the future actions of each institution, in addition to ensuring that the results are linked to the institutions future actions, this also generates a level of commitment in hierarchies or directors, regardless the processes or political wills.

Finally, since the institutionalisation of evaluation process in the country is not yet completed, there are no minimum standards or requirements that define the quality of evaluations or the skills of the evaluators

in charge of carrying them out; these issues represent the future challenges for Costa Rica and MIDEPLAN as the regulatory body in this regard.

While this is being achieved, MIDEPLAN has been concerned with generating documents, guidelines and guides that will allow it to cover basic aspects of evaluation in the institutions. An example of this is the MIDEPLAN (2012), which serves as a guide for institutions to carry out systematic and objective assessments of the design, development and results of evaluations in policies, programmes and projects. The management manual is an instrument that seeks to support results-based decision-making to improve public management.

The lack of standards is not only an issue for Costa Rica, and in view of its need, the Latin American and the Caribbean Evaluation, Systematisation and Monitoring Network (Red de Seguimiento, Evaluación y Sistematización de Latinoamérica y el Caribe, RELAC) have generated a document on evaluation standards for Latin America (2016), that seeks to become a reference framework that guides the evaluation processes in the countries of the region, allowing their adaptation to the social, political and economic contexts. This effort is highlighted as a step forward in the consolidation and institutionalisation of evaluations both in the region and in Costa Rica.

## DISSEMINATION IN SOCIETY AND ACCEPTANCE (SOCIAL SYSTEM)

### *Institutionalised Use of Evaluations by the Civil Society*

As seen in the previous sections, evaluations and their results are in the process of being consolidated as tools for parliamentary and political decision-making, reaching an initial level of awareness and training on the basic pillars that constitute them in some scenarios; this is why evaluation is not yet used to generate knowledge that influences decision-making at the political level. It is therefore no surprise that sectors outside the public administration do not even have knowledge of the evaluation processes, their implications and benefits, which leads to drastically reduced levels of use of evaluations and results.

What is said in the paragraph above applies to the majority of civil society, mainly if one looks at the levels and the capacity of use of evaluations by grassroots or community-based organisations. Therefore, it is not

possible to affirm that in Costa Rica evaluations and results are regular instruments for community decision-making; it is not even possible to state that the evaluations are greatly influencing action or decision-making within civil society organisations themselves.

The findings in the interviews show that civil society organisations as well as other sectors focus their efforts on providing accountability and meeting goals to donors on specific projects; and that there is still a lack of knowledge and training on evaluation and clear methodologies that allow for their participation, as well as concrete spaces and resources for it.

Nevertheless, it was observed that some sectors or non-governmental organisations (NGOs) such as the Central American Evaluation Association (Asociación Centroamericana de Evaluación, ACE) and the Evaluation and Monitoring Network of Costa Rica (Red de Evaluación y Monitoreo de Costa Rica, REDEVALCR) as well as individual consultants, had above-average professional capacities and technical tools, and it is this small sector of civil society that is responsible for carrying out in-depth research that could often meet the parameters of an evaluation, and that is aimed at the construction of alternative information or shadow reports of the situation of the country on a particular issue. These organisations also are following up and monitoring the actions of the state, mainly to verify that the official data are real, as a mechanism to confront official speeches, and that their advocacy work is producing concrete results.

In addition, there is a lack of spaces for civil society for participation in evaluation processes. Currently, public administration institutions do not have mechanisms or a methodology to articulate demands from civil society to request or participate in evaluation processes; MIDEPLAN has had some experiences of participatory evaluations that could be fertilising the ground where civil society can participate but there is not yet a clear tool in this regard. An example of this is that in the process of formulation and development of ANE the civil society has not received participation and that none of the 15 selected projects has been carried out under the participatory evaluation modality.

Therefore, the lack of knowledge and training on evaluations, the lack of political will to include the civil society as a participatory actor within the evaluation processes, and the lack of clarity on guidelines, parameters and methodologies make it difficult for civil society to use evaluations and results. This implies several challenges for the country, for its institutions

and mainly for the MIDEPLAN as the governing body of this matter, among which the need to train and educate on evaluation stands out.

In order for civil society to be able to articulate itself around the demands for evaluations, and mainly around the incorporation of evaluation results, it is necessary for the society to be informed and trained about the evaluation processes, their objectives, methodologies, and their impacts; without having any training on evaluation it is difficult for society to meaningfully participate in the processes, even if it counts on the required spaces to do it.

Similarly, education on evaluation should not only be provided so that society can give its opinion on policies, programmes and projects developed by the State, but can also be used as a tool to improve the very functioning of organisations, given that they are embedded in the culture of accountability.

The institutionalisation of a culture of evaluation does not only involve institutions of the public sector, it also implies the inclusion of civil society in all the processes that are occurring within the country. And their inclusion should not mean their confrontation with technical and professional evaluation processes, so it should be understood that initiatives and tools arise from civil society that offer particular and specific characteristics to the processes, and that therefore merit their entailment; such is the case of the social audits as a mechanism that allows the argumentative construction of this sector, and therefore their active participation in the processes.

### *Perception and Public Discussion of Evaluation and Results of Evaluation*

As stated in the previous section, evaluation tools and the use of evaluation results are still very limited in scope in the civil society sector, mainly due to a lack of knowledge and training in this area. Similarly, MIDEPLAN as the country's leading evaluation body, has focused its attention on permeating the evaluation processes internally and in other institutions that compose the public sector.

This is to be expected considering that the process has been carried out for a few years now, and that a national policy on the subject is being formulated. It would then be hoped that with the national evaluation policy, and some much more consolidated practices such as the ANE, the knowledge and experience that MIDEPLAN has acquired can

be transferred to civil society as a participatory actor with the capacity to contribute within the processes.

As for access to information and results of evaluations already carried out by the state, for example, evaluations within the framework of the ANE, the final reports are of public access and they are published in digital formats on the MIDEPLAN website. Additionally, some sectors have had the opportunity to participate in forums and conferences to update and provide information on the development of ANE and its results. However, information regarding forms or questionnaires used, or collected primary information is not of public access for the civil society.

Based on the interviews conducted, it was found that although the final reports of the projects are public, these do not count on the necessary level of promulgation to state that all public sector institutions or civil society organisations are aware of their results. Therefore, the fact that the reports are public is not a guarantee of knowledge and use within the institutions themselves, nor within sectors outside these institutions. The fact that information circulates, at least within public administration institutions, can be a factor that contributes positively to the level of knowledge in evaluation, as well as an opportunity for institutions to learn from real experiences.

### *The Demand of Evaluations in the Civil Society*

It is hoped that civil society will continue to have the opportunity to participate in spaces such as the national evaluation platform, and above all, in the construction of the national evaluation policy, since it is there, where the methodologies, models and times in which society can make a significant contribution to the evaluation processes are specified.

An evaluation process is known with the Ombudsman's office, the Ministry of Health, Social Security and other actors of civil society, which was a participatory evaluation, on the services of prevention and cancer care in the Valle la Estrella of the province of Limón financed and directed by FOCEVAL. This type of collaborative evaluation approaches has been little explored and there is no specific information on evaluations carried out by citizens, NGOs or private companies.

## PROFESSIONALISATION (PROFESSIONALISATION SYSTEM)

### *Academic Courses of Studies and Further Training*

In terms of academic training courses on evaluation, Costa Rica has a consolidated academic offer over time and it is highly recognised in the region; several disciplines and institutions have incorporated at least one evaluation component, almost always expressed in one subject at the undergraduate level, but the Master's degree from the UCR, and the CICAP are responsible for the professional training of evaluators in the country.

The Master's degree in evaluation of development programmes and projects of the UCR, initiated its work in 1995 under the modality of evaluation of social programmes and projects; years later, some modifications in the degree and curricular content of the postgraduate course allowed its transformation towards an approach of evaluation of development programmes and projects of development, and with it the opportunity to receive professionals from all the areas, which is why today it is considered a multidisciplinary professional master's degree. It is currently estimated that the postgraduate course has more than 100 graduate professionals and it is estimated that nearly 40% of them are working directly on evaluation (O. Villarreal, personal communication, March 1st, 2018).

As part of its evolution process, the Master's degree programme has been accompanied and advised by the Saarland University (UdS) in Germany, which has materialised concretely in the revision of the academic programme for year 2010, when the modifications that compose the current programme were made. To date, the master's degree is made up of four training blocks through which students are provided with conceptual, methodological, practical and personal tools for carrying out evaluations and communicating their results. In addition, it has a structured network of public sector organisations and institutions that benefit from evaluations carried out by postgraduate students.

The work of the UCR Master's degree is accompanied by the work of CICAP as a centre in charge of strengthening public administration organisations, through research, knowledge management, consulting and continuing education. Currently, CICAP represents the largest academic offer in continuing education in the country, by means of the direction of

seminars, short courses and technical programmes focused on the evaluation of policies, programmes and projects, and where nearly 700 public officials have been trained since 2005 to date.

CICAP has a specialised area of academic programmes on monitoring and evaluation, including the postgraduate seminar on impact evaluation; short courses on ‘evaluation for middle and upper management’, workshop on the construction of effect and impact indicators, financial evaluation of projects, and economic and social evaluation of projects; as well as technical programmes in Monitoring and Evaluation systems in the public sector of Costa Rica, and the manager’s programme on programme and project evaluation. In addition, the CICAP provides evaluation consultancy services, including two interventions in the ANE. Among the programmes offered by CICAP is the blended learning on evaluation. This programme comes from the Master of Evaluation Programme in Germany, which was first developed in Africa as a blended learning programme in 2017. It was initially proposed by Germany to Costa Rica as a training programme. After considering the proposal of Germany, CICAP showed its interest and from that moment began the process of organisation and coordination in order to be able to deliver the programme in Costa Rica. During 2018, GIZ paid for the translation of the modules from English to Spanish, CICAP reviewed the translated modules and executed two pilot tests and various adaptation efforts with the support of CEVAL of the UdS. The joint effort allowed this programme to be offered for the first time in 2018 for Costa Rica in blended learning format.

At the undergraduate, Bachelor’s levels, Costa Rica does not have a specialised training programme for evaluators, however, evaluation can be found in the form of a course in other scientific disciplines. This applies to the School of Social Work and the School of Public Administration, both of the UCR, which include courses on the formulation and evaluation of programmes and projects within their curricula. In addition, the International Centre for Economic Policy for Sustainable Development of the National University has training in evaluation in a very specific way within its academic offer.

CICAP has an offer of technical courses on evaluation of programmes and projects, financial evaluation and other short professional updating courses related to this topic. To receive this type of training, no specific knowledge is required although minimum requirements are requested



when they are public sector professionals and the requirements are eliminated when they are civil society or NGO collaborators, however, it is unusual for these people to be trained in this type of programmes.

### *Profession/Discipline*

At present, Costa Rica does not have specialised journals or bulletins in the subject of evaluation. Nevertheless, there are regional initiatives such as the Forum magazine, which was born as the first digital magazine specialised in evaluation of development policies, programmes and projects in the region; it launched its first issue in April, 2017.

The Forum magazine works also as a space for exchange ideas and experiences, in which Costa Rica participates through the support of the Evaluation Office of the Central American Bank for Economic Integration, the postgraduate programme in Evaluation of Development Programmes and Projects of the UCR, CICAP and the RELAC (Forum Magazine, 2018).

Another initiative is the production of research articles on the problems of Costa Rican public administration by CICAP. This compilation, although not entirely devoted to the evaluation topic, has included articles on it. The name of the journal is 'To Administer the Public'. This compilation of research articles emerged in 2013 and aims to systematise the interventions carried out by CICAP in the areas of consulting, advisory and training to provide input to decision makers and thus contribute to the analysis of opportunities within the institutions (CICAP, 2018).

The master's degree programme in evaluation of development programmes and projects has a series of professorships related to evaluation topics such as: evaluation approaches, planning and evaluation design, and meta evaluation, among others. This allows the discipline to be professionalised and improves the capacity of those involved in the learning of the evaluation, and also can contribute to explore intervention projects from the academy.

Regarding groups of evaluators, Costa Rica has three organisations, two with a regional approach and one with a national approach. Among the organisations with a regional approach are the RELAC and the ACE that currently has organisational problems that have hindered its work.

On the other hand, as a national initiative, there is the Evaluation and Monitoring Network of Costa Rica (REDEVALCR) that was born in 2017 as an association that seeks to contribute, spread and strengthen

evaluation in the country. Just one year after its creation REDEVALCR has made several contributions and has been present in important forums such as the national evaluation platform and the evaluation week in Latin America and the Caribbean. Its work focuses in three baselines, which are the strengthening of the culture of evaluation in the country, to offer spaces for diffusion and training on evaluation, and to generate spaces for dialogue and confluence of actors to influence the evaluation processes. (A.Bolaños & C.A. Montero, personal communication, May 09th, 2018).

The role of the organisations of evaluators such as REDEVALCR becomes important as a mechanism that helps to articulate the efforts that the Costa Rican government has been making for several years, with the opinions and contributions that can be given by civil society.

Finally, regarding the supply and demand for evaluators in the country, it could be affirmed that the Costa Rican labour market is small. Currently, the direct hiring of professionals in evaluation is handled in great part by the national evaluation agenda in charge of MIDEPLAN and some international organisations. The agenda, in addition to being the largest evaluation experience for the Costa Rican government, has become the perfect scenario for the participation of evaluators trained by the academic offer described above. It is important to mention that in previous periods before the agenda, the national market of evaluation was mostly occupied by international organisations interested in measuring their impact in the country, or by foreign evaluations contracted in Costa Rica where nationally trained evaluators were exported.

Despite its smallness, the market for evaluators is expanding. Costa Rica does not have a certification system for evaluators that allows to measure and ensure professional and technical skills, since it is considered that the university degree offered by the UCR postgraduate programme constitutes sufficient quality assurance. With respect to whether the country counts on an authority that has the role of an arbitrator in cases of unconformities on conducted evaluations, it can be said that it does not exist, therefore, common legal mechanisms for negotiating or terminating an evaluation contract are used; although these tools are applied to any type of contract under the Public Administration Law (Law N°6227 1978) they are the only formal mechanisms for resolving disputes in evaluation processes in the country.

### *Fulfilment of Norms and Quality Obligations*

As mentioned in previous paragraphs, Costa Rica has not yet defined standards on evaluation performance or quality. In view of this, organisations such as RELAC, and institutions such as MIDEPLAN have developed tools that, although they do not include any mandatory level, can serve to focus the work of evaluators in the country.

In 2016, RELAC, with the support of the FOCEVAL project presented a document on evaluation standards in Latin America, which mainly seeks to become a reference framework for the design, the hiring and the accomplishment of evaluations in Costa Rica as in the region. This document, although still in the enactment and dissemination phase, is an important step in helping to build national standards that define quality for evaluators and evaluations.

When CICAP has participated in evaluation processes of the ANE, it has formally requested respect for criteria such as autonomy and independence of the evaluator, mainly due to the contractor's need to control some aspects that go beyond his management. This exemplifies the need to have evaluation standards and comply with them. (E.O. Mora-Martínez, personal communication, Dec 4th, 2018).

On the other hand, and given the absence of national standards, evaluation clients have used the Terms of Reference (ToR) as mechanisms to measure quality and compliance within the evaluations. These ToR are formulated by the institution in charge of hiring the evaluation and are highly dependent on the type of evaluation to be carried out and its objectives. ToR are initially used to measure the abilities of evaluators at the time of recruitment, so it can be said that the institutions or organisations measure the training and experience of evaluators, but still remain short to measure the quality of the evaluations they carry out.

Given that the experiences in evaluation in the country have occurred with greater intensity in the recent years, institutions like MIDEPLAN have learned about the formulation of ToR as evaluations have progressed; this can be observed in the ToR for hiring evaluations within the NEA, with very ambitious terms, therefore not in accordance with the budget assigned for the evaluations.

Other institutions that have also hired evaluations directly have been technically supported by MIDEPLAN and FOCEVAL for the construction of ToR, which is a practice still being learned and consolidated in the country. In order to contribute to this process, MIDEPLAN has prepared

a guide to the ToR (MIDEPLAN, 2017), which aims to provide basic inputs on the technical and structural content that should be included in evaluations of public interventions. The guide shows not only that the ministry has acquired greater skills in this area, but also that is interested in consolidating this practice correctly within other institutions in the public sector.

## CONCLUSION

The institutionalisation of the evaluation culture in Costa Rica is not homogenous. While it is true that in recent years there has been greater regulation, more complex instruments have been developed, a large number of evaluations of projects of strategic interest have been carried out, the qualifications and training of the evaluating personnel as well of those who receive and participate in the management have increased, MIDEPLAN has improved its qualifications as the leading agency in the field with clear guidelines, instruments and trained personnel, there has been greater involvement of actors for the design and the implementation, and the use of results has been initiated, it can be concluded that there is no uniformity nor homogeneity in its application within the institutions.

By examining the culture of evaluation in public institutions more carefully, some dominant features of such culture can be identified, or at least outlined. Its most obvious signs which condition the activities and behaviours of the participants are: the achievement of goals and accountability, the central role played by the hierarchy from their conceptions, attitudes, values, traditions, practices and behaviours to the way they use evaluation; their privileged position of power and control within that context continue to be determinant in supporting evaluations and in using their results; and finally there are asymmetric relationships between the evaluator and the evaluated subject, since the evaluator continues to control the entire evaluation process and the evaluation subject is submitted to the system.

In spite of this, the perception about the importance of evaluation for decision-making has changed thanks to MIDEPLAN's work with the support of FOCEVAL and the built platform, and with the academy's work on training and capacity building.

The sustained work of these actors with the inclusion of civil society gives optimism that the future will be better to strengthen the process already underway, to institutionalise a culture of evaluation in the country.

## LIST OF ABBREVIATIONS

ACE	Asociación Centroamericana de Evaluación/Central American Evaluation Association
AES	Área de Evaluación y Seguimiento/Evaluation and Monitoring Area
ANE	Agenda Nacional de Evaluación/National Evaluation Agenda
CGR	Contraloría General de la República/Comptroller General's Office of the Republic.
CICAP	Centro de Investigación y Capacitación en Administración Pública/Centre for Research and Training in Public Administration
FOCEVAL	Proyecto de Fortalecimiento de Capacidades en Evaluación/Project for Capacity Building in Evaluation
GIZ	German Corporation for International Cooperation GmbH
MIDEPLAN	Ministerio de Planificación y Política Económica/Ministry of National Planning and Economic Policy
NGO	Non-governmental organisation
PND	Plan Nacional de Desarrollo/National Development Plan
POI	Plan Operativo Institucional/Institutional Operational Plans
REDEVALCR	Red de Evaluación y Monitoreo de Costa Rica/Evaluation and Monitoring Network of Costa Rica
ReLAC	Red de Seguimiento, Evaluación y Sistematización de Latinoamérica y el Caribe/Latin America and the Caribbean Evaluation, Systematisation and Monitoring Network
SINE	Sistema Nacional de Evaluación/National Evaluation System
ToR	Terms of Reference
UCR	University of Costa Rica

## REFERENCES

- CICAP. (2018). *Regresa administrar lo público con resultados de investigación y casos de estudio sobre gestión pública*. Centro de Investigación y Capacitación en Administración Pública—Universidad de Costa Rica. <http://www.cicap.ucr.ac.cr/web/administrar-lo-publico/>. Accessed on 14 November 2019.
- Comisión de Coordinación Interinstitucional. (2011, November). *Decreto Ejecutivo N°36901*. Sistema Costarricense de Información Jurídica, Procuraduría General de la República. [http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm\\_texto\\_completo.aspx?param1=NRTC¶m2=1&nValor1=1&nValor2=72051&nValor3=87712&strTipM=TC&lResultado=2&nValor4=1&strSelect=sel](http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?param1=NRTC¶m2=1&nValor1=1&nValor2=72051&nValor3=87712&strTipM=TC&lResultado=2&nValor4=1&strSelect=sel). Accessed on 13 July 2018.
- Costa Rica. 2001. *Asamblea Legislativa de Costa Rica, Budget and Financial Administration Law*, No. 8131, San José.
- Fonseca. A. (2001). *El Sistema Nacional de Evaluación: un instrumento para la toma de decisiones del gobierno de Costa Rica*. XV Concurso de Ensayos del CLAD ‘Control y Evaluación del Desempeño Gubernamental’. Caracas, Venezuela.
- Forum Magazine. (2018). <https://www.evalforum.com/>. Accessed on 13 July 2018.
- García, E., & Ugalde, K. (2015). *Panorama de los sistemas nacionales de monitoreo y evaluación en América Latina*. Centro de Investigación y Docencia Económica—Centro CLEAR para América Latina. Ciudad de México: México.
- Law of Financial Administration. (2006, January). *Decreto Ejecutivo N°32988*. Sistema Costarricense de Información Jurídica, Procuraduría General de la República.
- Law of National Planning. (1974, May). *Ley de Planificación Nacional N°5525*. Sistema Costarricense de Información Jurídica, Procuraduría General de la República. [http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm\\_texto\\_completo.aspx?param1=NRTC¶m2=1&nValor1=1&nValor2=72051&nValor3=87712&strTipM=TC&lResultado=2&nValor4=1&strSelect=sel](http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?param1=NRTC¶m2=1&nValor1=1&nValor2=72051&nValor3=87712&strTipM=TC&lResultado=2&nValor4=1&strSelect=sel). Accessed on 13 July 2018.
- Law of Public Administration. (1978, May). *Ley General de la Administración Pública N°6227*. Sistema Costarricense de Información Jurídica, Procuraduría General de la República. [http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm\\_texto\\_completo.aspx?param1=NRTC¶m2=1&nValor1=1&nValor2=72051&nValor3=87712&strTipM=TC&lResultado=2&nValor4=1&strSelect=sel](http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?param1=NRTC¶m2=1&nValor1=1&nValor2=72051&nValor3=87712&strTipM=TC&lResultado=2&nValor4=1&strSelect=sel). Accessed on 25 July 2018.
- Law N°8003. (2000, June). *Reforma del artículo 11 de la Constitución Política*. Sistema Costarricense de Información Jurídica, Procuraduría General de la República. <http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/>

- nrm\_texto\_completo.aspx?param1=NRTC&nValor1=1&nValor2=29804&nValor3=31496&strTipM=TC. Accessed on 13 July 2018.
- Law N°8131. (2001, September). *Ley de la Administración Financiera de la República y Presupuestos Públicos*. Sistema Costarricense de Información Jurídica, Procuraduría General de la República. [http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm\\_texto\\_completo.aspx?param1=NRTC&m2=1&nValor1=1&nValor2=72051&nValor3=87712&strTipM=TC&lResultado=2&nValor4=1&strSelect=sel](http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?param1=NRTC&m2=1&nValor1=1&nValor2=72051&nValor3=87712&strTipM=TC&lResultado=2&nValor4=1&strSelect=sel). Accessed on 13 July 2018.
- MIDEPLAN. (1994, May). *Decreto Ejecutivo N°23323*. Sistema Costarricense de Información Jurídica, Procuraduría General de la República. [http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm\\_texto\\_completo.aspx?param1=NRTC&m2=1&nValor1=1&nValor2=72051&nValor3=87712&strTipM=TC&lResultado=2&nValor4=1&strSelect=sel](http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?param1=NRTC&m2=1&nValor1=1&nValor2=72051&nValor3=87712&strTipM=TC&lResultado=2&nValor4=1&strSelect=sel). Accessed on 1 July 2018.
- MIDEPLAN. (2010). *Plan Nacional de Desarrollo 2011–2014* ‘María Teresa Obregón Zamora’. Gobierno de Costa Rica. <https://www.mideplan.go.cr/plan-nacional-desarrollo-2011-2014>. Accessed on 13 July 2018.
- MIDEPLAN. (2012). *Manual gerencial para el diseño y ejecución de evaluaciones estratégicas de gobierno*. Programa Fortalecimiento de las Capacidades en Evaluación en Centroamérica (FOCEVAL).
- MIDEPLAN. (2014). *Plan Nacional de Desarrollo 2015–2018* ‘Alberto Cañas Escalante’. Gobierno de Costa Rica. <https://www.mideplan.go.cr/Plan-Nacional-Desarrollo-2015-2018>. Accessed on 13 July 2018.
- MIDEPLAN. (2017). *Guía de términos de referencia. Orientaciones para su elaboración: estructura y contenido*. Área de Evaluación y Seguimiento. Costa Rica.
- MIDEPLAN. (2019). *Sistema Nacional de Evaluación*. <https://www.mideplan.go.cr/sistema-nacional-evaluacion>. Accessed on 13 July 2018.
- Political Constitution. (1949). *Costa Rican legal information system*. Office of the attorney general of the republic. [http://www.pgrweb.go.cr/scij/busqueda/normativa/normas/nrm\\_texto\\_completo.aspx?param1=NRTC&nValor1=1&nValor2=871&strTipM=TC](http://www.pgrweb.go.cr/scij/busqueda/normativa/normas/nrm_texto_completo.aspx?param1=NRTC&nValor1=1&nValor2=871&strTipM=TC). Accessed on 13 July 2018.
- MIDEPLAN. (n.d.). *Estudio exploratorio: estado de situación del Sistema Nacional de Evaluación* (SINE).
- SINE. (2010, January). *Decreto Ejecutivo N°35755*. Sistema Costarricense de Información Jurídica, Procuraduría General de la República. [http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm\\_texto\\_completo.aspx?param1=NRTC&nValor1=1&nValor2=67309&nValor3=79671&strTipM=TC](http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?param1=NRTC&nValor1=1&nValor2=67309&nValor3=79671&strTipM=TC). Accessed on 13 July 2018.

## LIST OF INTERVIEWS

- José Chaves—Associate Auditor, Comptroller General of the Republic.
- Olman Villareal—Director, Master's Degree in Evaluation of Development Programmes and Projects.
- Florita Azofeifa—Director, Monitoring and Evaluation area, Ministry of National Planning and Economic Policy.
- María Lourdes Jaén—Unit Coordinator, Budget Management Unit, Ministry of Finance.
- Andrea Meneses—Technical Manager, FOCEVAL, Costa Rica.
- Jeannette Carrillo—Director, Dissemination and Enactment, *Defensoría de los Habitantes*.
- Miriam Valverde and Ghiselle Rodríguez—Institutional Planning Unit, Executive Secretary to the Agricultural-sector Planning, Ministry of Agriculture and Livestock farming.
- Ailhyn Bolaños and Carlos Alberto Montero—Evaluation and Monitoring Network of Costa Rica.
- Rosibel Vargas—Director, Strategic Institutional Planning and Development, Ministry of Health.
- Juan Carlos Morales—Representative of civil society organisations to the National Evaluation Platform.
- Rita Flores—Representative of the Permanent Forum, civil society organisations.
- Víctor Quesada y Lucía López—Institutional Planning Unit, Costa Rican Institute of Tourism.
- Ana Luisa Guzmán—Former coordinator, Latin American Monitoring and Evaluation Network.
- Mario Céspedes—Representative of civil society to the National Evaluation Platform, Civil Society Organisations.
- Edith Paniagua—Director, Library Services, Legislative Assembly.
- Esteban O. Mora-Martínez—Evaluator and Research-Innovation Coordinator of CICAP.





## CHAPTER 9

---

# Evaluation in Ecuador

*Cinthia Josette Arévalo Gross  
and Lourdes Cumandá Montesdeoca Espín*

### GENERAL COUNTRY OVERVIEW

In order to have a clear understanding of the situation of evaluation in Ecuador, it is important to consider the historical and political processes that have set the stage for the development and institutionalisation of evaluation. Ecuador is a small developing country located in South America. The total population of Ecuador was 16.6 million people in 2017. Around 21.5% of the population was below the poverty line, and life expectancy at

---

Josette Arévalo Gross is affiliated with the Office of Evaluation and Oversight (OVE) of the Inter-American Development Bank (IDB). The opinions expressed in this article represent those of the authors and not necessarily those of IDB Group.

---

C. J. Arévalo Gross (✉)  
Inter-American Development Bank (IDB), Washington, DC, USA

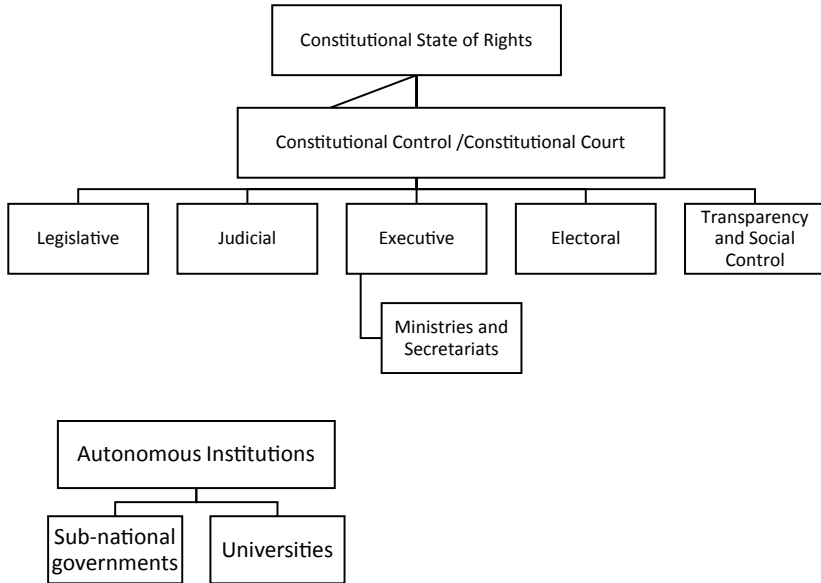
L. C. Montesdeoca Espín  
Facultad Latinoamericana de Ciencias Sociales, Quito, Ecuador  
e-mail: [lmontesdeoca@flacso.edu.ec](mailto:lmontesdeoca@flacso.edu.ec)

birth was 76.3 years in 2016 (World Bank, 2018a). Although the completion rate for primary school was close to 98% in 2014, the completion rate for secondary school was only close to 65% (Unesco Institute of Statistics, 2018). Furthermore, less than 7% of adults have a bachelor or higher degree (INEC, 2010), and there are only 84 Ph.D. or research professionals per one million inhabitants (World Bank, 2018b). These statistics provide a general idea of some of the educational and social conditions of the population that influence the country's capacities for the institutionalisation of evaluation.

Regarding the sociopolitical settings and administrative system, Ecuador is a representative democratic republic; the government signed the last version of its Constitution in 2008 after popular approval through a national referendum. This Constitution replaced the one signed ten years before, and it is the twentieth version that has been in place since 1830. According to the current constitution, administratively, Ecuador is divided into five powers: executive, legislative, judicial, electoral, and the transparency and social control function (Asamblea Nacional del Ecuador, 2008; see Fig. 9.1). Regarding the geographic organisation, there are 24 provinces, 221 cantons and 1225 towns; from the latter, 811 are rural towns whereas 414 are urban towns. Nevertheless, there is a high concentration of administrative activities and population in the two largest cities: Quito (the administrative capital city) and Guayaquil (the biggest and the most commercial city in the country). Finally, it is important to highlight that Ecuador has suffered frequent political instability throughout its history, with significant civil, political and territorial confrontations. The most notable social and political unrest in recent history was 1997–2007, with a rate of ten presidents in ten years (Acosta, 2012; Paz y Miño Cepeda et al., 2007).

*Why are these country features important to the study of evaluation institutionalisation in Ecuador?*

The socioeconomic conditions of the country, especially the educational attainment levels, are important in the institutionalisation of policy evaluation since they determine the demand of evaluations from society, and the supply of professionals that are trained in evaluation. International public debt has been an important restriction in public finances throughout the country's history. International Development Agencies have been the main source of finance, planning and evaluation for large-scale infrastructure projects. These characteristics, in addition to the



**Fig. 9.1** Ecuadorian state functions and institutions (SENPLADES, 2012)

political instability and the educational attainment, are important predictors that explain most of the outcomes in public policies (Acosta, 2012). Thus, apart from the socioeconomic conditions, the political conditions and the way the administrative system is organised are important determinants for the level of institutionalisation of policy evaluation.

This chapter focuses on three main elements of the process of institutionalisation of evaluation in Ecuador. First, this chapter discusses institutional structures and processes including evaluation laws and regulations, evaluation practice and use of evaluations. Second, it describes Ecuador's situation regarding the dissemination of evaluation and acceptance by society. Third, a description of the professionalisation of evaluation of the country is provided. Lastly, some conclusions and an outlook of the future of institutionalisation of evaluation in Ecuador are presented. The chapter was written based on literature review, a review of legal documents, and on the responses from twelve evaluation experts from the public sector, NGOs, international organisations and private sector.

## POLITICAL SYSTEM: INSTITUTIONAL STRUCTURES AND PROCESSES

This section provides a review of Ecuadorian evaluation regulations, evaluation practice and use of evaluations. In general, the emergence of evaluation processes in Ecuador goes in hand with the institutionalisation of planning processes. The newest constitution, signed in 2008, includes specific articles about evaluation (which are then operationalised in laws and regulations) and describes principles of planning and evaluation as a mandatory requirement for the public sector. Regarding evaluation practice, this section describes the level of evaluation diffusion, the most common types of evaluations that are conducted, and the levels of independency of evaluation institutes and divisions in Ecuador. Lastly, in order to understand the use of evaluation, this section describes the main purposes of evaluation and quality assurance mechanisms.

### *Evaluation Regulations*

#### *Brief Historical Review of Planning and Evaluation*

Literature related to a historical approach of evaluation in Ecuador is very scarce or almost non-existent. Nevertheless, given that evaluation is linked to planning processes, we approach this part through a review of the history of planning in Ecuador.

Up to 1954, there is no official documented or institutional evidence of central planning or evaluation in Ecuador. Before 1954, governments basically based their budgets and decisions on projects or laws established on pressures of economic, military and religious interest groups. During the first half of the twentieth century, there were two remarkable periods related to central planning. The first one was the Juliana Revolution (1925–1931), a period where institutions like the Central Bank of Ecuador were created under the advice of the Kemmerer mission. The second one was the ‘*Plan Estrada*’ which was banned by the congress in 1934 (Acosta, 2012).

After these two periods, the first national planning institution created in Ecuador was JUNAPLA (*Junta Nacional de Planificación y Coordinación Económica*, National Planning and Economic Coordination Board), which was established on May 28th, 1954 through an executive decree. In 1979, almost at the end of the military government (1972–1979), JUNAPLA was replaced by another institution

called CONADE (*Consejo Nacional de Desarrollo*, National Development Council). CONADE included affiliated institutions such as INEC (*Instituto Nacional de Estadística y Censos*, National Institute of Statistics and Census), the National Fund of Pre-Investment and the *Consejo Nacional de Ciencia y Tecnología* (National Science and Technology Council).

In 1998 CONADE was replaced by ODEPLAN (*Oficina de Planificación*, Planning Office). In 2004 the SENPLADES (*Secretaría Nacional de Planificación y Desarrollo*, National Planning and Development Secretary), was created by Executive Decree No. 1372. This new institution was the result of the merge of ODEPLAN and *Secretaría de Diálogo Social y Planificación*. Since then, and up until mid-2019, SENPLADES was in charge of national planning and evaluation, with an M&E division within its organisational structure. In addition to these planning institutions, SODEM (*Secretaría Nacional de los Objetivos de Desarrollo del Milenio*, National Secretariat of the Millennium Development Goals) also had an important role in planning as well as evaluating public policies, focused mainly on the social sector. It was created under the international agenda of Millennium Development Goals. SODEM was merged to SENPLADES by Executive Decree No. 103 of February 22th, 2007 (SENPLADES, 2018).

Regarding M&E, most of the plans that were carried out by these planning institutions were monitored in short terms (mainly on a yearly basis), and mostly by external agencies like the IMF (1980s). The first national and technical attempt of an evaluation was in 1992 through the Strategic and Situational Planning methodology; then the vision changed from annual monitoring to a continuous monitoring of activities and processes as well as to ‘permanent and systematic’ evaluations. Yet, in practice, evaluations were implemented only for two years: 1994 and 1996 with CONADE on charge. It was not until ODEPLAN when systematically and institutionalised control, monitoring and evaluating processes were implemented, for that purpose a methodological guide and IT systems were created (Ullauri Enriquez, 2002).

### *Formalisation of Evaluation in the Constitution, Laws and Regulations*

This subsection presents a brief overview of the main ways in which evaluation has been formalised in the constitution, and most important laws and regulations. Due to the length and scope limitations, this section only brushes upon major legal instruments, and does not pretend to provide

an exhaustive legal review. There are several other sector-specific laws and regulations, in which evaluation is at least considered; some of these are listed in Appendix 7.1.

### *Constitution*

Contrary to previous Ecuadorian constitutions, the current political constitution of Ecuador, signed in 2008, includes not only the principles of planning and evaluation as a mandatory requirement for the public sector, but it also includes some specific articles about evaluation (Asamblea Nacional del Ecuador, 2008):

- Article 227: Says that the principles of public administration are: efficacy, efficiency, quality, hierarchy, de-concentration, decentralisation, coordination, participation, planning, transparency and evaluation.
- Article 255: the *Sistema Nacional de Planificación* is established as a technical secretary with dependency on the Presidency office, also the local governments and social organisations must be under this system.
- Article 280: establishes the National Development Plan called *Plan Nacional de Desarrollo* (National Development Plan, PND) which is not only an administrative tool for public policy that has four years of length, but also it is a policy agenda for every new president and authorities.
- Article 346: indicates that an autonomous institution in charge of internal and external educational evaluation/assessment will be created in order to promote quality education.

### *Laws*

Since 2008, other legal instruments have been approved or modified<sup>1</sup> in order to institutionalise planning and evaluation. Most of these legal instruments are of general application and other are specific for certain sectors such as education. One of the main legal instruments regarding evaluation is the *Código Orgánico de Planificación y Finanzas Públicas* (Organic Code of Planning and Public Finance, COPFP), in which article number 26 establishes the responsibilities of SENPLADES. This article indicates that SENPLADES is in charge of carrying out the monitoring and evaluation of the National Development Plan and its

<sup>1</sup> According to the new political constitution.

instruments; promoting and carrying out relevant studies for national planning; and proposing technical inputs for the consideration of the National Planning Council. Regarding the National Decentralised System of Participatory Planning (*Sistema Nacional Descentralizado de Planificación Participativa*), the COPFP indicates that public administration has to be goal-oriented and results-oriented, and it needs to take into consideration tangible and intangible impacts. Since planning and evaluation go hand-in-hand, the COPFP provides an important framework regarding SENPLADES' responsibilities, in conducting evaluations related to the PND and provides guidance for public institutions.

### *Regulations*

The regulation of the COPFP (Reglamento al COPFP) provides more detailed and specific information about the role of SENPLADES in M&E. For instance, article 54 of this regulation indicates that: it is the duty and attribution of SENPLADES, "to lead the national subsystem of monitoring and evaluating of public interventions to achieve the objectives and goals of the PND" as well as planning, directing and accompanying the design and implementation of methodologies for monitoring and evaluation of public interventions (Presidencia de la República, 2014). In addition, article 8 indicates that the National Council of Planning is in charge of approving the Annual Evaluation Plan (*Plan Anual de Evaluación*, PAEV).

Nevertheless, it is important to mention that SENPLADES was dissolved by Executive Decree No. 732 of March 14th, 2019. Some of SENPLADES' former responsibilities were transferred to the Ministry of Finance, and others (i.e. planning) are now part of the newly created National Secretariat '*Planifica Ecuador*', which is under the command of the Presidency. At the moment this chapter was written, there was still some uncertainty regarding how this will affect the before-mentioned laws and regulations.

### *Formalisation of Evaluation in the Legislative Branch*

There is no specific formalisation of evaluation in the legislative branch. However, in the same fashion as with public institutions from the executive branch, each member of the National Assembly has to present an annual accountability report, which sometimes is regarded as an evaluative tool. This report is made according to the '*Ley de Participación Ciudadana*' which establishes that all popular elected authorities are

obliged to render an account regarding the fulfillment of their work plans to the citizens (CPCCS, 2014). Nevertheless, as one of the interviewees mentioned, it is a common practice that not all individual congressmen present a unique report; the Assembly's different commissions present their reports as a summary of the activities of all its members. If assembly members do not present their annual reports, the *Consejo de Participación Ciudadana y Control Social* (Council of Citizen Participation and Social Control, CPCCS) reports this to the National Control/Auditing Agency for its investigation.

The National Assembly can also request the results of the national evaluation plan from the Council of National Planning and could require evaluation reports from other branches. In addition to the legislative function, there is a constitutional mandate demanding the supervision of "(...) the acts of the executive, electoral and transparency and social control functions, and the other organs of the public power, and require the servants and public servants the information they consider necessary" (Asamblea Nacional del Ecuador, 2008, Art. 120 § 9). Furthermore, the President must present an annual report to the National Assembly (Asamblea Nacional del Ecuador, 2008, Art. 120 § 4).

### *Evaluation Practice*

#### *Diffusion of Evaluation Practice*

In order to understand the diffusion of evaluation practice in Ecuador, it is important to understand the scope and frequency in which evaluation is conducted by actors from different sectors of society. The interviews with evaluation experts from the public, civil society and legislative sectors and international organisations indicate that policy and programme evaluations tend to be of national coverage and are concentrated on certain sectors. The heterogeneity in the diffusion of evaluation practice among sectors appears to be linked to the financing sources and the affinity of the authorities in charge. Specifically, the increasing amount of social sector programme evaluations tends to be related to international organisations/donor's demands of evaluation or their accompaniment in the evaluation process. Still, most interviewees indicate that there is a limited reach, scope and capacity of evaluation practice in all sectors of society.

The central government, along with international organisations, and academia are the predominant actors that conduct policy and programme evaluations. As mentioned before, SENPLADES is in charge of evaluating



the PND and of conducting the evaluations listed in the PAEV approved by the National Council of Planning (Presidencia de la República, 2014, Art. 8 of Reglamento COFP). These evaluations are mainly concerned with emblematic programmes with national coverage. According to the Chief of the Evaluation Division of SENPLADES, the criteria to choose the policies to be evaluated in the PAEV include: policies related to low performing PND indicators, policies related to PND, number of beneficiaries, budget, coverage and prioritisation by the Presidency. Due to the limited resources available, only a few programmes or policies are included in the PAEV.

Regarding the frequency of evaluation, interviewed experts agree that apart from the PAEV, there is no institutionalised frequency in which actors from different sectors conduct evaluations. As a former public sector evaluation authority indicates, ‘There is no specific rhythm or frequency. There is no evaluation culture’. Furthermore, due to budget restraints and capacity, only a limited number of policies are evaluated annually. Moreover, although SENPLADES requires that all agencies and ministries include an evaluation component in their investment project proposals when they apply for budget approval, some experts agree that this is merely a formal requirement that does not translate into the execution of project evaluations. Often times the proposed ‘evaluations’ in investment project proposals could be better described as basic monitoring activities with limited methodological descriptions, and do not have a specific frequency. Due to the volume of investment projects executed in the public sector each year, there are important limitations to verify whether the evaluations are conducted and even more difficulties for quality assurance.

The frequency in which policies and central government programmes are evaluated is erratic and mostly depends on the willingness and interest of the authorities in charge. Some experts agree that public sector authorities are more willing to evaluate policies and programmes at the beginning of their management period, since the results will be attributable to previous administrations. In addition, interviewed experts indicate that the frequency in which evaluations are conducted is lower and, in some cases, non-existent in the legislative branch and at the subnational level (local governments).

### *Most Common Types of Evaluations*

The different types of evaluation can be distinguished by who conducts the evaluation (internal or external), the timing (ex-ante, medias res, or expost), and the focus (inputs, outputs, process, results and impact).

Although evaluation practice is heterogeneous in Ecuador, interviews with evaluation experts and a qualitative study conducted by Villarreal et al. (2018) shed some light on the most common types of evaluation conducted in the country. According to several interviewed experts, the most common types of evaluations that are conducted in the public sector tend to be external evaluations and output/process evaluations. Nevertheless, there is a prevalence of monitoring processes which are often mistakenly labelled as ‘evaluations’. In general, evaluations tend to focus more often on outputs rather than on results/outcomes. Yet, it is important to mention that the only recurrent evaluations in the public sector are conducted by SENPLADES, which generally consist in external impact/results evaluations of other ministries’ programmes/policies. In addition, there is a growing interest and movement towards conducting impact evaluations.

Interviewed experts cited four main reasons why external evaluations are more common than internal evaluations. The first reason is related to the influence and/or requirements from international organisations and donors. Either international organisations explicitly require external evaluations or other times they conduct the external evaluations themselves. The second reason given by some interviewees is that external evaluations are more common because they tend to be considered more credible and legitimate. Thirdly, ministries and institutions that provide services often need to focus more on service provision which limits the time to conduct internal evaluations (which is also why there are more internal monitoring processes than evaluations). Lastly, poor evaluation capacities and scarce resources that can be explicitly dedicated to M&E activities within most public institutions at the central level (and even worse in local governments), limit the amount of internal evaluations that can be conducted. Nevertheless, it is important to mention that institutionalised continuous monitoring processes are conducted by teams within public institutions which report the progress of their strategic indicators to SENPLADES through the *Gobierno por Resultados* (Government by Results system, GPR). Therefore, although external evaluations are often the most common type of evaluations conducted in Ecuador, internal

monitoring processes are prevalent throughout public institutions. Villarreal et al. (2018) present similar results from document reviews and surveys conducted to 17 public institutions and international organisations located in Ecuador: evaluations are mostly conducted by a mix of internal and external evaluators depending on the project (53%), there is a clear confusion between monitoring and evaluation, and most institutions (67%) conduct monitoring processes on a regular basis.

Regarding the focus of evaluations, there are several reasons that explain the prevalence of evaluations (or rather monitoring activities) that are focused on process and outputs rather than on results or impact. Interviewees explained that monitoring activities are more common than results/impact evaluations due to limitations in the sizes of M&E teams, and also on limitations on the knowledge of evaluation methods and concepts among M&E teams within most public institutions. These perceptions are corroborated by Villarreal et al. (2018) study which finds that monitoring processes are conducted more regularly than evaluations, yet the study's results also indicate that results evaluations are more frequent than impact and process evaluations. Moreover, among institutions that do conduct evaluations, results evaluations are more common than impact evaluations.

In addition, interviewed experts indicate there is a strong preference towards measuring 'what can be measured easily', so M&E teams and programme managers within public institutions tend to choose activity and output indicators rather than results/impact indicators. As a former public sector evaluation specialist indicated,

"(...) there is a very poor evaluation culture, or a negative view of evaluation... so when we wanted to help with strategic planning, we sought in an utopian manner that, for example, the ministry of health would evaluate itself by coverage, and not by the number of vaccines... but they chose the indicator of the number of vaccines because it was easier for them to report the information that way instead of thinking about the problem in terms of coverage".

This is also related to the strong preference of having more immediate information for accountability purposes. Therefore, M&E teams tend to focus on monitoring more than evaluations.

Although evaluations focused on outputs and outcomes are the most common, impact evaluations are starting to gain space and recognition.

There are even some specialists that argue that the only type of good quality evaluation that is conducted in Ecuador are impact evaluations. This type of evaluation is more prevalent on large-scale programmes funded by international organisations, but is also the preferred type of evaluation conducted by researchers that wish to publish in international academic journals. Moreover, they mainly evaluate the results of education or cash transfer programmes.

### *Evaluation Institutes and Divisions, and Their Levels of Independency*

Institutionalised evaluation instances are not common in Ecuador, and the ones that exist vary in their characteristics: coverage, sectoral scope, and independency. Although there is no national independent evaluation institution in Ecuador that covers all sectors, SENPLADES has a specific division that is in charge of public policy M&E and of conducting the evaluations listed in the PAEV. Although the division is not independent nor autonomous of SENPLADES, it is independent of the institutions that are evaluated. Moreover, the education sector has two national independent evaluation institutions (INEVAL and CACES). There is also an increasing number of divisions or units within ministries that conduct M&E with varying and lesser degrees of independence. Additionally, the National Control/Auditing Agency (*Contraloría General del Estado*) is in charge of auditing all public sector institutions, and the four sectoral Superintendences are in charge of controlling and auditing the institutions under their supervision.

According to SENPLADES' Organisational Process Statute (SENPLADES, 2016, Art. 10 #1.2.2.4.3), the Directorate of Public Policy Evaluation (Dirección de Evaluación de Políticas Públicas) within the Monitoring and Evaluation Division (Subsecretaría de Seguimiento y Evaluación) was in charge of designing methodologies and procedures for policy, programme and project evaluation. In addition, it was in charge of conducting the evaluations listed in the PAEV, and of providing technical assistance to other public sector institutions regarding policy, programme and project evaluation methodologies. Lastly, it is in charge of providing recommendations to help accomplish project/programmes' goals. Thus, the unit within SENPLADES conducted evaluations that are independent from the ministries that own the programmes. As mentioned before, since 2019 some of these responsibilities are now part of the newly created

National Secretariat '*Planifica Ecuador*', which is under the command of the Presidency.

Within the education sector there are two independent and autonomous institutions in charge of evaluation of K-12 education and higher education: INEVAL and CACES. INEVAL (*Instituto Nacional de Evaluación Educativa*, National Institute of Educational Assessment) is an autonomous public institution created, on paper, by the 2008 Constitution. Article 346 indicates that a public and autonomous institution will be created to promote quality education through integral internal and external evaluation/assessment. INEVAL has a degree of independency since it is not financially nor administratively dependent on the Ministry of Education, yet the president of its board is the Minister of Education. The board members, as well as the Executive Director of INEVAL, have to meet high academic and professional qualifications, mandated by law (LOEI article 72).

INEVAL, which was officially created in 2012, periodically assesses students in third, sixth, ninth and twelfth grades (*Ser Estudiante* and *Ser Bachiller* assessments), conducts an assessment of public-school teachers (*Ser Maestro*), and is in charge of conducting international large-scale assessments such as PISA. The results of these assessments are important tools for policymakers in the educational arena. Although INEVAL mainly focuses on educational assessment (i.e. educational achievement testing and teacher assessment), it also has an educational research unit with the capacity to conduct programme evaluations.

The Council of Higher Education Quality Assurance (CACES) oversees the assessment of Bachelor's degrees and programmes' educational achievement (Asamblea Nacional del Ecuador, 2018). Before CACES, there were several different types of institutions in charge of higher education evaluation/assessment (CONUEP, CONEA and CEAACES). In 1989 CONUEP conducted the first institutional evaluation of universities and higher education institutions. In 2002, CONEA started to operationalise self-evaluation, external evaluation and accreditation processes (Villavicencio, 2008). From 2008 to 2011 universities' institutional evaluations were conducted by request. The 2010 Organic Law of Higher Education (LOES) created the Council of Evaluation, Accreditation and Quality Assurance of Higher Education (CEAACES), which worked from 2011 to 2017. The 2018 reforms of LOES changed the scope of CEAACES attributions, and renamed the council: CACES, removing the evaluation and accreditation portion from its name. The reformed higher

education law (LOES) indicates that CACES is in charge of regulation, planning and coordinating the higher education quality assurance system, and that it is an independent institution (Asamblea LOES, Art. 171). In addition, according to LOES' article 173, CACES is in charge to regulate higher education institutional self-evaluation, will conduct external evaluation, will be in charge of accreditation, and will support internal quality assurance processes.

Regarding M&E within public institutions, during the past decade, the executive branch has institutionalised the creation of monitoring and evaluation units within central government public institutions. To a lesser extent, these units have been created as research units within ministries. Across the public sector, these units are usually situated either as stand-alone divisions (*Direcciones*) within a Strategic Planning Coordination, or as units within Strategic Planning Divisions. Yet, Villarreal et al. (2018) indicate that, although the majority of the 17 public institutions and international agencies that participated in the survey do conduct M&E, almost half of them indicated that there are no specialised M&E units in their organisational structures. Moreover, two out of ten indicated that M&E is conducted across different units (Villarreal et al., 2018). It is important to mention that although the M&E units tend to be independent in the organisational structure, their independence level tends to be questioned because they are part of the same institution.

### *Use of Evaluations*

Although there is widespread acknowledgement of the importance of evaluation among project managers and administrative regarding how evaluation results should become important inputs to learn and to improve policy interventions, there is still no generalised evaluative culture to complete the cycle of planning, implementing, monitoring and evaluating. In addition, due to limited financial resources and professional capacities, reporting indicators and/or filling forms are often regarded as 'evaluations', when these activities could only, in the best case, be considered monitoring activities. In order to understand the use of evaluations in Ecuador, this section discusses evaluation purposes and use, followed by a discussion about quality assurance processes.

### *Evaluation Purposes and Use*

As Chelimsky and Shadish (1997) indicate, the use that can be expected of an evaluation is conditioned by the purpose of evaluation. Therefore, to understand evaluation use, it is important to first understand the purposes of evaluation. Patton indicates that apart from the three primary purposes of evaluation cited in previous editions of his 'Utilisation Focused Evaluation' book (rendering judgements, facilitating improvements and generating knowledge) there are three additional purposes: accountability, monitoring, and development (Patton, 2008).

The majority of the expert interviewees indicate that the most common purposes for evaluation are rendering judgement, accountability and monitoring. In addition, the evaluations that are conducted are basically focused on administrative outputs in order to justify the use of public resources and meet legal requirements such as '*rendición de cuentas*', which is a mandatory annual requirement for all public institutions. Nevertheless, the documents produced to meet these legal requirements are far from being rigorous evaluations.

Most of the experts agree that rendering judgement could be the main political motivation to require and implement an exhaustive evaluation. There is the perception that the majority of bureaucrats and common people do not know the utility of the evaluation beyond possible retaliations or political costs. These evaluations are often used to justify the existence or the need to expand/cut-back programmes. There are also few incentives to evaluate for planning and knowledge purposes. The accountability purpose, which is predominant among evaluations, is technically and legally established (mostly in administrative and financial terms). There is a national public institution (*Contraloría General del Estado*) in charge of auditing the use of public resources and which holds other institutions accountable. Therefore, public institutions often conduct evaluations with the purpose of being accountable to the public, but also to meet legal requirements. Thus, the use given to these evaluations is limited to the purpose of accountability.

Regarding the monitoring purpose, as mentioned earlier, according to Ecuadorian legislation, public institutions have a unit within their organisational structures in charge of planning, monitoring and evaluating. In general, these units focus on monitoring activities, which are used to report to authorities. In the executive branch for instance, since 2012 a management by results tool called GPR was put in place. Public institutions must report monitoring results in the GPR tool. GPR has four

levels of monitoring: the first level, with the president office in charge it has a national coverage and comprises cross-cutting programmes as well as emblematic (or large national) projects, the second level is the sectoral programmes in charge of each minister or the main authority of the institutions that execute public policies; the third level is the coordination unit in each institution, in charge of institutional monitoring and finally the executing units of each project or public policy in charge of their final implementation and monitoring (Grandes Villamarín, 2016). Thus, due to administrative and legal requirements, the monitoring purpose is widespread among public institutions. This purpose conditions the use of the results, which are often limited to produce required reports.

Lastly, the purposes related to facilitating improvements, development and generating knowledge are not as common. Professors or researchers are some of the few stakeholders that are usually interested in these evaluation purposes. According to interviewees, the majority of citizens and programme managers do not tend to consider these three possible purposes due to lack of knowledge or resources.

Regarding sectors that conduct and use evaluations, there was a *boom* of evaluations in the social sector mainly since 2000, with special emphasis on the education sector or cash transfers. According to some interviews, other sectors of ‘good performance’ in evaluation use are the security, transportation and the economic sectors. On the other hand, according to most interviewees, the environmental sector and the majority of subnational governments have been historically the least evaluated.

There are different groups of people who use evaluations. The main group who uses evaluations are the project or programme managers in order to show their success—if the evaluation is positive—or to avoid responsibilities otherwise (rendering judgement, monitoring and accountability purposes), but also in a lesser extent to have information about whether programmes should continue or need to be improved (rendering judgement and facilitating improvements purposes). The second group of interest is the academia and/or researchers with the goal of conducting studies and publishing (generating knowledge purposes). Lastly, the third group of interest is the administrative or control personnel (mostly for accountability and monitoring purposes). Nevertheless, there is a consensus between interviewees that politicians and civil citizens, in general do not know about evaluations nor use evaluations.



### *Quality Assurance*

From the legal point of view, the regulation of COPFP (Article 60) indicates that SENPLADES is the national and legal institution that will regulate the minimum technical parameters of quality and its procedures for reporting and dissemination of the evaluations of public sector institutions. According to SENPLADES, the quality of evaluations is based on a permanent accompaniment process of design, implementation and execution from the SENPLADES' Directorate of Public Policy Evaluation, which follows the established parameters on the 'Guide for managing public intervention evaluation processes' (*Guía para Gestionar los Procesos de Evaluación de Intervenciones Públicas* in Spanish).

However, there is a certain agreement among expert interviewees that only a few people have knowledge about those regulations. Moreover, the people who have knowledge about the regulations consider them as formal requirements, rather than as technical guides. They said that 'administrators and project managers only fill forms'. As one of the interviewees indicated:

"There is no guarantee (of quality), and because of that, there are a lot of poor quality evaluations; their errors are later discovered by replications or evaluations carried out by academia or by (independent) consultants, or studies financed by international organisations".

This generalised perception has a negative impact on public sector administration as well as on contractors who work for public institutions. On the extreme, they perceive those requirements as a 'waste of time and resources, without added value'. Therefore, although there are guidelines for quality assurance, administrators, managers and general citizens have not internalised their importance. Consequently, the guidelines are under-utilised, and reporting and studies are seen as a bureaucratic requirement instead of technical approaches to learn and improve public policies.

### DISSEMINATION/ACCEPTANCE BY THE SOCIETY

As pointed out in the previous section, there is no positive perception of the implementation and use of evaluation, mainly due to the lack of knowledge about what an evaluation process entails and what it is useful for. This situation is reinforced by two extreme situations: on one hand, there is no evaluation information at all from some public institutions, and

on the other hand there could be too much information about evaluation results that is not systematic nor usable for decision making. Moreover, there is an important gap related to public access of information and communication. As interviewees mentioned, the Ecuadorian society's lack of knowledge or interest about evaluation may be used by politicians who would want to exaggerate any of the results (good or bad), as long as the results benefit them electorally. Furthermore, interviewees mentioned that politicians tend to not even consider evaluation as an opportunity to learn and improve public policies.

In order to understand evaluation dissemination and acceptance by society, this section will describe the use of evaluations by civil society, the public perception and discussion of evaluation and evaluation findings, and civil society's demand of evaluation.

### *Institutionalised Use of Evaluations by Civil Society*

According to most interviewees, the general perception is that in Ecuador there is no institutionalised use of the evaluations, neither in the public sector, private sector, nor by civil society. Civil society rarely discusses or demands policy/programme evaluations. According to interviewees, this is due in part to the lack of capabilities and knowledge about evaluation, and also due to the fact that when evaluations are conducted, their results tend to be of internal use and knowledge. Therefore, evaluation results and evaluative processes are generally unknown and, in many cases, not promoted.

Although there is no institutionalised use of evaluations for electoral processes, referendums or community-based policymaking, several interviewees point out two mechanisms of information exchange between citizens and the public sector that may inform some decisions and perceptions about programme/policy performance: the dialogue with citizens and the accountability system. The former is a new mechanism implemented by the government in the country (as of May 2017), focused on collecting the best ideas to improve public policies. The accountability system is called '*rendición de cuentas*', which is the annual mandatory report that all elected officials have to present. These two mechanisms may (but not very often) integrate some evaluation results, but the use of evaluation results is not institutionalised nor generally demanded by civil society.

Among NGOs, there are some important examples of initiatives that represent important progress towards institutionalised use of evaluations. First, Grupo FARO, an Ecuadorian non-profit think tank, has spear-headed an initiative called ‘Ecuador Decide’ with a subcomponent: ‘*Del Dicho al Hecho*’ (‘From Said to Done’). ‘*Del Dicho al Hecho*’ is a project that aims to reduce the information gap between government actions and what citizens know. This project’s first phase consisted in evaluating the government’s campaign promises after 100 days, six months and one year in power. The evaluation was conducted using input, process and results indicators that were focused on 6 areas: technical education, early childhood education, employment, entrepreneurship, fight against corruption and the reconstruction of Manabí province after the earthquake. This project has been the first of its kind, and represents an important advance towards a more institutionalised use of evaluation by civil society (Ecuador Decide, 2019).

Another initiative led by Grupo FARO is ‘*Datos Ciudadanos para los ODS*’ (Citizen Data for the Sustainable Development Goals, SDGs). This initiative seeks to improve the monitoring and implementation of the SDGs through empowerment, information and support to national and local articulation (Grupo FARO, 2019). Lastly, Grupo FARO and Fundación Futuro Latinoamericano lead the ‘*ODS Territorio Ecuador*’ project funded by the European Union. This project aims to improve the livelihood of the population and reaching the SDGs through national and local policies and the strengthening of civil society for the implementation and monitoring processes. The project includes creating one national and five local citizen observatories that will be in charge of monitoring the progress of reaching the SDGs, and also building citizen capacities of monitoring the SDGs (ODS Territorio Ecuador, 2019).

Although there is no widespread use of evaluations among NGOs and the private sector, there have also been some important initiatives that promote evaluation use in the private sector. *Pacto Global Ecuador* is a private initiative inserted in the United Nations Global Compact Network and the Ecuadorian Consortium for Social Responsibility. This initiative is governed by 10 universally accepted principles in four thematic areas: human rights, labour standards, environment and anti-corruption. Specifically, this consortium seeks to channel actions in support of the 2030 Agenda SDGs. Currently the network has 80 active members, including private and public companies, together with local governments,

academia and NGOs. The initiative holds an annual event to disseminate the best practices of organisations around the 2030 Agenda at the international level. Participating institutions must: (1) Incorporate the principles of Pacto Global as part of its strategy and (2) communicate annually the progress in the implementation of the principles through reports. The latter stands out as an institutionalised monitoring strategy (Pacto GlobalRed Ecuador, 2019).

### *Public Perception and Discussion of Evaluation and Evaluation Findings*

As pointed out in the previous section, the majority of citizens and even most public servants do not know about evaluation, and evaluations findings are rarely discussed. There is only a reduced and very selected group of professionals that know about evaluation and evaluative methods, and who are interested in discussing them. Nevertheless, there is usually no distinction regarding the different types of evaluation that exist. Some interviewees indicate that “international organisations, including the private sector, are aware of this, but they often remain in process evaluations, and do not go further to improve interventions. They do not comply with rigorous quality. Many times it is merely a descriptive analysis and evaluation is confused with only showing if the budget was implemented”.

In addition, there is a generalised perception that evaluation results are not used in public discussions, there are usually no lessons learned after programme implementation, nor recommendations for future programme improvement. Therefore, new similar programmes start almost from zero, instead of learning from past interventions. This situation often happens because information about monitoring and evaluating programmes ‘is lost’, especially when authorities (and technical teams) are changed and do not leave programme documents for new authorities. In addition, evaluations results are often only known by authorities. Depending on whether the results are positive or negative, some studies may be archived as reserved, or in the best scenario placed in a virtual library which is only accessed and used by a few people. From the evaluations that are published, very few of them are comprehensive evaluations or impact evaluations, and they are generally published without peer reviews. Although there are no dedicated journals for communicating evaluation in Ecuador, evaluation articles have been published in journals dedicated to other

disciplines, such as economics and social sciences, and university journals. For instance, impact evaluations have sometimes been published in the *National Institute of Statistics and Census (INEC)* journal.

Despite the concerns about the low level of knowledge, diffusion and use of evaluation results, it is important to mention that discussions about evaluation have been increasing in the past few years. In addition, some public institutions have started to conduct research and evaluation in accordance with current regulations. Furthermore, there is political will, there are more funds to do more technical researches and evaluations, and there is growing interest in evaluation among higher education institutions. For instance, some Ecuadorian universities like Flacso, ESPOL, EPN and Pontifica Universidad Católica del Ecuador (PUCE) have organised some meetings and workshops about evaluation methods (most of them focused on the education sector), and some events have also been organised by civil society organisations. One example of this is the ‘evidence week’, an event organised by civil society organisations in late 2017 where international and national experts were invited to present their studies and new methodologies. In 2019, several groups (PUCE, EvalYouth Ecuador, EPN and EvalEC) organised various events and workshops for the 2019 CLEAR gLOCAL Evaluation Week.

### *Civil Society’s Demand of Evaluations*

Most citizens do not know what a comprehensive and rigorous evaluation is, nor how to distinguish a good quality evaluation from a bad quality evaluation. According to some interviewees, citizens tend to associate the word ‘evaluation’ with the results from an evaluation, so the focus is on whether ‘the project is finished’ instead of what should be covered and the quality of the project. Citizens’ low educational attainment and the very limited knowledge about evaluation affect the demand for rigorous evaluations. Nevertheless, there are some initiatives from civil society organisations that demand them (i.e. Esquel Foundation, Fundación Futuro Latinoamericano, and Grupo FARO). Grupo FARO, for instance, mapped in 2018 the actors that could demand an evaluation. The findings indicated that although private enterprises implement programmes, they usually do not demand evaluations. When private enterprises conduct evaluations, they tend to be process evaluations. There is also some demand of evaluation from civil society, especially when funds come from

elsewhere, and from enterprises that are committed to a social cause. As an interviewee from the civil society indicates:

B enterprises, as they are committed to a cause, want to know the impacts they have on society. This is a new niche. Individual citizens also demand evaluations, but it depends on the organisation. Sometimes when they are organised in groups. For example: the observatory of the SDGs where they ask to empower themselves on the objectives, but there is a demand for accountability.

### PROFESSIONALISATION (SYSTEM OF PROFESSIONALISATION)

In Ecuador, at the time when this chapter was written, there were no institutions or graduate programmes specialised in evaluation, and because of that (among other reasons), there are few evaluation professionals. Most of the few evaluation professionals working in Ecuador have gained their knowledge by their own means (e.g. through international/online courses or seminars, or by participating in evaluation processes) or by studying abroad. Moreover, most of them are not evaluation experts; rather they are generally experts in other fields but have learned some evaluation concepts and methods. In order to comprehend the professionalisation of evaluation in Ecuador, this section discusses three important issues: academic courses and further training, a description of the evaluation profession/discipline, and compliance to standards and quality obligations.

#### *Academic Study Courses and Further Training*

At the time when this chapter was written, there were no specific academic evaluation programmes in Ecuador at the undergraduate or postgraduate levels (Rodríguez-Bilella, 2018). Although there are no recurring evaluation programmes, there are some sporadic and isolated courses that are offered in programmes from other professional fields (e.g. within Masters of Economics programmes like FLACSO). There are also some courses that are offered sporadically by international organisations or by demand from public institutions for specific purposes. Due to the lack of in-person specialised courses, people interested in learning about evaluation often search for and rely on online options. According to interviewees, some

of the well-known virtual courses are the ones offered by J-PAL or Ed-X courses for evaluation of social programmes, and those designed and implemented by the IDB or the World Bank.

According to some higher educational institutions, the lack of supply of specialised evaluation academic programmes is in part due to the lack of demand of this kind of programmes. For example, the Catholic University of Ecuador offered in 2015 a Master degree in Economics with a major in '*Public Policy*', which contained quantitative evaluations courses, but the programme was closed and later reformulated due to an insufficient number of students that enrolled in the programme. The Catholic University of Ecuador has opened the Master's programme again; but now in coordination with the public Central University of Ecuador and Grenoble University from France (PUCE, 2019). Furthermore, the Instituto de Altos Estudios Nacionales (IAEN), started working on a new Master's programme about '*Evaluation Techniques*'. IAEN expected the programme to be operational by the end of 2020.

In July and August 2017, EvalYouth Ecuador—a network of young and emerging evaluators created in 2017—conducted an online survey aimed to understand the situation and challenges that Ecuadorian evaluators face. Although the sample is small ( $N = 29$ ), the survey sheds some light on the perceptions of evaluators and other professionals interested in the field. Most people who answered the survey were 26–30 years old, and either worked in the public sector or academia. The results of this survey show that although most respondents indicate that there are some evaluation courses available, eight out ten of respondents perceive that there is not enough quantity or quality of courses (EvalYouth Ecuador, 2017).

The increasing demand and interest in evaluation, evidenced for example by the creation of EvalYouth Ecuador, has a good outlook for the future. For instance, an interviewee mentioned that there is a project of building a postgraduate programme specialised in evaluation within the Ecuadorian Catholic University. According to the interviewee, the project was in its implementation phase, where future national trainers were being trained by international experts.

### *Profession/Discipline*

There are few civil society organisations, networks and NGOs that are interested in evaluation aspects in Ecuador. Yet, there are three evaluation networks in Ecuador. Two of them are informal voluntary evaluation networks: *Red de Evaluadores del Ecuador* and *EvalYouth Ecuador*. The third one, Sociedad Ecuatoriana de Evaluación (SEEval) is a formal and legally established evaluation network, created in 2019. According to its president, the *Red de Evaluadores del Ecuador* is an independent network, whose main objective is the interaction between its members. It has been working for almost 15 years since its creation and does not have external financing sources. It is self-financed with the contributions and collaboration of its members. It was part of the Latin American evaluation network (ReLAC) and many of the events carried out have been financed with funds from international organisations and with personal contributions. According to the network's representative, the network does not have a legal status yet due to bureaucratic obstacles that require certain formal requirements such as the RUC (*Registro Único de Contribuyentes* in Spanish) and keeping financial accounts. According to the network's president, another limitation for the expansion of the network has been the time that volunteers can provide to the network. Due to lack of funding, she said that, at the moment, there is no record of the members and the formalisation and governance structure of the network is in process (there are no governance processes established yet). From the records of events and email list, the network estimates that it has about 40 members of all ages.

EvalYouth Ecuador was created in 2017 as an Ecuadorian network of young and emerging evaluators. This national chapter is in line with the global EvalYouth initiative, whose goals are to promote Young and Emerging Evaluators (YEE) to become competent, experienced and well-networked professionals who contribute to evaluation capacity at national, regional and international levels, and to promote the inclusion of Youth and Young People (YYP), including young women, in evaluations conducted at the national, regional and international levels (EvalYouth nd). EvalYouth Ecuador has a governance structure with an executive committee composed of eight members, which have to be elected every year: Two co-chairs, and six members in charge of three task forces: (1) Strengthening the presence of YEE in the governance of Voluntary Organisations of Professional Evaluators (VOPE); (2)



Creating and implementing a mentoring programme on evaluation for YEE; (3) Improving the capacities of YEE through face-to-face and virtual conferences.

Since its creation in June 2017 and up until 2019, EvalYouth Ecuador had organised six regular workshops with its members. In addition, EvalYouth participated as a co-organiser of the Evidence Week 2017, which took place in October 2017. During this week, EvalYouth organised several workshops related to impact evaluations of public programmes implemented in Ecuador. Although this network is new, EvalYouth Ecuador has already created some important partnerships. For instance, EvalYouth Ecuador has partnered with Foceval (a regional project that aims to develop evaluation capacities and that is financed German Ministry of Development and Economic Cooperation) to foster knowledge exchange with young evaluators in Costa Rica. As part of this partnership, members of EvalYouth Ecuador have participated in trainings and other events related to evaluation both in Ecuador and Costa Rica. EvalYouth has also organised capacity building events for the past three years during CLEAR's Evaluation Week.

### *Compliance to Standards and Quality Obligations*

Professional organisations rarely ask their members to follow rules or guiding principles related to quality. In the case of SENPLADES, following the guidelines is mandatory, but, there is a gap between regulations and practice. Moreover, according to several interviewees, the guidelines of SENPLADES are usually not known by most evaluators. Some institutions and NGOs do not demand a certain quality of evaluation and/or compliance with standards. Moreover, interviewees indicated that, in many cases, it is difficult to evaluate the quality of evaluations because people in general do not have the capacities to do so. The requirements of compliance with standards sometimes comes from NGOs and international organisations, who may require quality assessments for the approval of products and payments.

## CONCLUSIONS AND OUTLOOK

This chapter described important elements of the process of institutionalisation of evaluation in Ecuador, but it is important to indicate that the levels of institutionalisation vary across executive and legislative powers,

and subnational governments, the latter being the least developed. The first important aspect to understand the institutionalisation of evaluation are the institutional structures and processes. In the Ecuadorian context, it is important to note that the emergence and development of evaluation processes goes in hand with the institutionalisation of planning processes that have grown in scope and strength in the past decade, which have also been bolstered and materialised in the newest Constitution, and subsequent laws and regulations. The results of the interviews with experts indicate that, although evaluation practice is heterogeneous, there is a limited reach, scope and capacity of evaluation practice in all sectors of society. In addition, the frequency in which policies and programmes are evaluated is erratic and usually depends on the interest and willingness of the authorities in charge. Regarding evaluation use, most evaluations are used for rendering judgement, accountability and monitoring purposes, and are often conducted only to meet accountability reporting (*rendición de cuentas*) or monitoring requirements.

The major challenges for the institutionalisation of evaluation in Ecuador are the dissemination of evaluation, acceptance by society, and the professionalisation of evaluation. Interviewees agree that there is no institutionalised use of the evaluations, neither in the public sector, private sector, nor by civil society. Since civil society rarely discusses or demands policy/programme evaluations, public sector authorities are not pressured to produce more evaluations nor to improve the quality of the evaluations. This issue of poor knowledge about what evaluation is and what 'good quality' evaluation is, goes in hand with the incipient professionalisation of evaluation. At the time this chapter was written, there were no specialised higher education evaluation programmes in Ecuador, which is a limiting factor since evaluation professionals need to seek training options abroad, train themselves, attend sporadic available courses or attend virtual courses. In addition, voluntary networks of evaluators are still concentrated in the capital city, and the reach and knowledge about their existence is still limited. Yet, the increasing demand and interest in evaluation have a good outlook for the future, depending in part on the political stability of the country. Evaluation networks are growing, undergraduate and graduate programmes are starting to include evaluation courses, and even a specialised evaluation course is being designed. Even if it is only for accountability purposes, evaluation is starting to be demanded. Now one of the biggest challenges is to build and strengthen a culture of evaluation.

## LIST OF ABBREVIATIONS

CACES	Higher Education Quality Assurance Council ( <i>Consejo de Aseguramiento de la Calidad de la Educación Superior</i> )
CONADE	National Development Council ( <i>Consejo Nacional de Desarrollo</i> )
COPFP	Organic Code of Planning and Public Finance ( <i>Código Orgánico de Planificación y Finanzas Públicas</i> )
CPCCS	Council of Citizen Participation and Social Control ( <i>Consejo de Participación Ciudadana y Control Social</i> )
GPR	Government by Results ( <i>Gobierno por Resultados</i> )
INEC	National Institute of Statistics and Census ( <i>Instituto Nacional de Estadística y Censos</i> )
INEVAL	National Institute of Educational Assessment ( <i>Instituto Nacional de Evaluación Educativa</i> )
JUNAPLA	National Planning and Economic Coordination Board ( <i>Junta Nacional de Planificación y Coordinación Económica</i> )
ODEPLAN	Planning Office ( <i>Oficina de Planificación</i> )
PAEV	Annual Evaluation Plan ( <i>Plan Anual de Evaluaciones</i> )
PND	National Development Plan ( <i>Plan Nacional de Desarrollo</i> )
SDGs	Sustainable Development Goals ( <i>Objetivos de Desarrollo Sostenible</i> )
SENPLADES	National Planning and Development Secretary ( <i>Secretaría Nacional de Planificación y Desarrollo</i> )
SODEM	National Secretariat of the Millennium Development Goals ( <i>Secretaría Nacional de los Objetivos de Desarrollo del Milenio</i> )

## APPENDIX

Other laws and regulations related to the planning and evaluation in Ecuador.

1. CODIFICACION DEL CODIGO DEL TRABAJO
2. CODIGO DE LA NIÑEZ Y ADOLESCENCIA

3. CODIGO ORGANICO DE LA FUNCION JUDICIAL
4. CODIGO ORGANICO GENERAL DE PROCESOS
5. CODIGO ORGANICO INTEGRAL PENAL
6. CODIGO ORGANICO DE ORGANIZACION TERRITORIAL
7. CODIGO ORGANICO DE PLANIFICACION Y FINANZAS PUBLICAS
8. ESTATUTO REGIMEN JURIDICO ADMINISTRATIVO FUNCION EJECUTIVA
9. LEY ORGANICA DE SERVICIO CIVIL Y CARRERA ADMINISTRATIVA
10. LEY ORGANICA DE EDUCACION SUPERIOR, LOES
11. LEY ORGANICA DE PARTICIPACION CIUDADANA

## REFERENCES

- Acosta A. (2012). *Breve Historia Económica del Ecuador* (3rd ed.). Corporación Editora Nacional.
- Asamblea Nacional del Ecuador. (2008). *Constitución Política de la República del Ecuador*. Montecristi, Ecuador.
- Asamblea Nacional del Ecuador. (2018). *Ley Orgánica de Educación Superior. Registro Oficial Suplemento 298 de 12 October 2010*.
- Chelimsky, E., & Shadish, W. R. (Eds.). (1997). *Evaluation for the 21st century: A handbook*. Sage.
- Consejo de Participación Ciudadana y Control Social—CPCCS. (2014). *Guía Especializada de Rendición de Cuentas para la Función Legislativa*. <http://www.cpccs.gob.ec/wp-content/uploads/2017/01/7-GUIA-ESPECIALIZADA-DE-RENDICION-DE-CUENTAS-PARA-LA-FUNCION-LEGISLATIVA.pdf>. Accessed September 20, 2018.
- Ecuador Decide. (2019). *Que es Del Dicho al Hecho? Grupo Faro*. <https://deldichoalhecho.ecuador-decide.org/>. Accessed June 31, 2019.
- EvalYouth Ecuador. (2017). *Diagnóstico Jóvenes y Emergentes Evaluadores en Ecuador (Online Survey)*.
- EvalYouth. (nd). *What does EvalYouth want to achieve?*. <https://www.evalpartners.org/evalyouth>. Accessed September 13, 2018.
- Grandes Villamarín, R. D. (2016). *Diagnóstico y propuesta de mejora de la herramienta Gobierno por Resultados (GPR) en las instituciones de la administración pública dependientes de la función ejecutiva del gobierno ecuatoriano 2011–2014 caso: Agencia de Regulación y Control Hidrocarburífero (ARCH)*. Master's thesis.

- Grupo FARO. (2019). *Datos Ciudadanos para los ODS*. <https://grupofaro.org/areas-de-trabajo/datos-ciudadanos-para-los-ods/>. Accessed June 31, 2019.
- Instituto Nacional de Estadística y Censos (INEC). (2010). *El Censo Informa: Educación. Resultados del Censo de población y vivienda en el Ecuador*. [http://www.ecuadorencifras.gob.ec/wp-content/descargas/Presenciones/capitulo\\_educacion\\_censo\\_poblacion\\_vivienda.pdf](http://www.ecuadorencifras.gob.ec/wp-content/descargas/Presenciones/capitulo_educacion_censo_poblacion_vivienda.pdf). Accessed August 25, 2018.
- ODS Territorio Ecuador. (2019). *Sobre la Iniciativa*. <https://odsteritorioecuador.ec/antecedentes/>. Accessed June 30, 2019.
- Pacto Global Red Ecuador. (2019). Una iniciativa global con impacto local. *UN Global Compact*. <http://www.pactoglobal-ecuador.org/unete-a-la-nueva-era-de-negocios-responsables>. Accessed October 18, 2020.
- Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Sage.
- Paz y Miño Cepeda, J. J., Aguilar, P. D., & de la Torre Muñoz, C. (Eds.). (2007). *Asamblea Constituyente y Economía. Constituciones en Ecuador*. Taller de Historia Económica.
- Pontificia Universidad Católica del Ecuador. (PUCE). (2019). *Maestría en Economía, mención Políticas Públicas*. <https://www.puce.edu.ec/portal/carreras/maestria-en-economia-mencion-politicas-publicas/>. Accessed October 18, 2020.
- Presidencia de la República. (2014). *Reglamento Código Orgánico de Planificación y Finanzas Públicas*. Decreto Ejecutivo 489. Registro Oficial Suplemento 383 de November 26, 2014. Last modified on October 25, 2016.
- Rodríguez-Bilella, P. (2018). Más de cuarenta posgrados en evaluación en América Latina y el Caribe. *Al Bordo del Caos*. <https://albordedelcaos.com/2018/05/17/mas-de-cuarenta-posgrados-en-evaluacion-en-america-latina-y-el-caribe/>. Accessed December 9, 2018.
- SENPLADES. (2012). *Reforma Democrática del Estado*. <http://www.planificacion.gob.ec/wp-content/uploads/downloads/2012/08/Reforma-Democr%C3%A1tica-del-Estado.pdf>. Accessed August 8, 2018.
- SENPLADES. (2016). *Estatuto Orgánico de Gestión Organizacional por Procesos de la Secretaría Nacional de Planificación y Desarrollo*. Acuerdo No. SNPD-044-2016. <https://www.planificacion.gob.ec/wp-content/uploads/downloads/2016/11/Registro-Oficial-Estatuto-Senplades-755.pdf>. Accessed August 8, 2018.
- SENPLADES. (2018). *Plan Estratégico Institucional 2018–2021*. Retrieved from: <https://www.planificacion.gob.ec/wp-content/uploads/downloads/2018/08/Plan-Estrategico-Senplades-2018-2021.pdf>.
- Ullauri Enriquez, C. (2002). Nueva Visión de la Planificación y el Desarrollo en el Ecuador, el anteproyecto de Ley. *Instituto de Altos Estudios Nacionales*. Master's thesis.

- Unesco Institute of Statistics. (2018). *UIS. Stat.* <http://data.uis.unesco.org/>. Accessed July 27, 2018.
- Villarreal, A., Castells, P., & Castro, A. (2018). Evaluación de programas y políticas públicas en Ecuador: Oportunidades y desafíos. *ValorAgregado*, 10, 43–74. <http://valoragregado.ec/>. Accessed January 6, 2019.
- Villavicencio, A. (2008). Los nuevos retos del aseguramiento de la calidad de la educación superior en el Ecuador. In *Seminarios Internacionales por la Calidad de la Educación Superior. Seminario Internacional de la Evaluación y Acreditación. Aportes para pensar la Educación Superior* (pp. 191–204). Secretar.
- World Bank. (2018a). *World Bank Data.* <https://data.worldbank.org/country/ecuador>. Accessed July 27, 2018.
- World Bank. (2018b). *World Bank Data.* <https://datos.bancomundial.org/indicador/SP.POP.SCIE.RD.P6>. Accessed August 25, 2018.



## Evaluation in Mexico

*Janett Salvador Martínez and Jaqueline Meza Urías*

### INSTITUTIONAL STRUCTURES AND PROCESSES

#### *Federal Legal Framework*

Mexico is a federal republic, with a representative and democratic government. The federation is composed of 32 federative entities which have democratically elected autonomous governments.

Since 2007, performance of evaluation of social programmes has been institutionalised by the Mexican Federal Government. The early history took place in 1995 when the evaluation was inserted by Presidential Decree in a programme called ‘*PROGRESA*,’ which was considered one of the main activities carried out by the National Programme Coordination for measuring its impact and provide feedback on its government operation (DOF, 1997).

Subsequently, from the legislative branch, programmes were subject to operating rules included in the Federation’s Expenditure Budget to be evaluated externally to know the impact of their results.

---

J. Salvador Martínez (✉) · J. Meza Urías  
Mexican Association of Evaluators (ACEVAL), Mexico City, Mexico  
e-mail: [Janett.salvador@c-evalua.mx](mailto:Janett.salvador@c-evalua.mx)

J. Meza Urías  
e-mail: [jaqueline.meza@comunidad.unam.mx](mailto:jaqueline.meza@comunidad.unam.mx)

In 2005, under the General Law of Social Development (LGDS), the National Council for Evaluation of Social Development Policy (CONEVAL) was created “to review social programmes periodically to comply with social development policy goals and objectives to amend, modify, add, redirect or suspend them partially or entirely” (Cámara de Diputados del H. Congreso de la Unión, 2004, p. 16). The following year in 2006, the Federal Budget and Treasure Act was passed, and the Performance Evaluation System was created (SED) to evaluate the impact of programmes to measure the results of evaluations based on their budget (Cámara de Diputados del H. Congreso de la Unión, 2006, p. 67).

In 2007, the executive branch issued the General Guidelines for Federal Programmes Evaluation of the Federal Public Administration, with the purpose to “regulate the evaluation of federal programmes, prepare the matrix of results indicators and monitor systems, as well as establish ministries strategic objectives of the Public Federal Administration” (CONEVAL & SHCP, 2019, p. 1). Subsequently, Article 34 of the Political Constitution of the United Mexican States was amended to institutionalise the evaluation of federal, state and local public resources.

Therefore, the national evaluation policy was established within a legal framework, which supports ministries’ and government agencies’ duties as well as different systems to collect information on compliance results and national budget programmes (Pp) being the CONEVAL and Ministry of Financy and Public Credit (SHCP) the ministries in charge of regulating the evaluation and follow up activities of Pp through SED and the System of Aspects Susceptible for Improvement (ASM) respectively.

The Mexican evaluation legal framework is summarised in Table 10.1.

**Table 10.1** Mexican evaluation legal framework (own development)

<i>Government</i>	<i>Legal Framework</i>
Federal	<ul style="list-style-type: none"> <li>• Political Constitution of United Mexican States (CPEUM)</li> <li>• Federal Budget and Treasure Act (Articles 25, 85, 110 and 111)</li> <li>• General Law of Government Accountability (Articles 54 and 80)</li> <li>• Tax Coordination Law (Article 49)</li> <li>• General Law of Social Development (Articles 72 to 80)</li> <li>• General Law on Government Accounting</li> </ul>
State and Local	<ul style="list-style-type: none"> <li>• General guidelines for federal Pp evaluation of the Public Administration</li> <li>• Aspects Susceptible for Improvement tool derived from reports and Pps evaluation</li> </ul>



### *Sectoral Federal Legal Framework*

As mentioned, from 1994 to 2007, the Mexican State was building the regulatory framework of the monitoring and evaluation system (M&E) aimed at achieving results, aware that the evaluation “plays an important role in the modernisation of the government apparatus and in the impact of their investments” (Kliksberg & Rivera, 2007, p. 130), in addition to that it allows to articulate “the management with the results and allows to make transparent the actions of the State, enabling the social control of the people on the commitments assumed” (Kliksberg & Rivera, 2007, p. 131).

As determined by article 134 of the CPEUM, all programmes and budgets of the agencies and entities that involve economic resources of the federal government must be evaluated to assess their performance and to know the effectiveness, efficiency and honesty with which they are executed “to meet the objectives to which they are intended” (Eval-uare, 2019, n.p.). However, when reviewing the existing literature on evaluation in Mexico, most of the evaluations carried out correspond to the social, educational and environmental sectors. This could be due to the fact that evaluation priorities in Mexico have been oriented towards accountability, policy improvement and, ultimately, decision-making, since the evaluation has been linked to social rights by what is important to know, in a preponderant way, the diagnosis of social problems because, as an article published by the *SEE-Valuation Magazine* points out, the thematic priorities of impact evaluations in most Latin American countries are: food sovereignty and fight against hunger, education, social and labour protection and health (Larrú, 2010).

The fact that the evaluation has a greater development in the social, educational and environmental sectors is also due to the fact that, after the financial crisis in Mexico between the 1980s and the 1990s, the evaluation of government performance made it possible to assess the economic efficiency of the programmes and personnel to reduce the disproportionate and unnecessary public expenditure made during the previous years, and begin to reestablish the financial balance in the federal public administration.

In this environment of reorganisation of public finances it was logical that the evaluation arose in the field of social policies (and especially in education [...]), which consumes huge public resources, and which has often been used politically to gain support popular, without the great public spending being able to prove that it has favourably transformed

the social conditions of its recipients, and modified the most harmful and unacceptable general social situations. (Cardozo, 2006, p. 10)

Those three sectors of national development rely on a legal framework and ministries in charge of coordinating and regulating public policy evaluation in different areas, as shown in Table 10.2.

**Table 10.2** National sectors with a greater improvement in evaluation (own development)

<i>Sector</i>	<i>Evaluation Scope</i>	<i>Legal Framework and Regulatory Organisations</i>
Social Development	Poverty Measurement Social Development Programmes Evaluation	<ul style="list-style-type: none"><li>• Political Constitution of United Mexican States (Article 26)</li><li>• National Council for Social Programme Evaluation (CONEVAL)</li></ul>
Education	<ul style="list-style-type: none"><li>• Performance Evaluation of the National Education System and Teachers Evaluation</li></ul>	<ul style="list-style-type: none"><li>• National Institute for Education Evaluation Act</li><li>• General Act of Professional Teachers Service</li><li>• National Institute for Education Evaluation (in process of repeal)</li><li>• National Institute for the Evaluation of Education (INEE)</li><li>• National Institute of Ecology and Climate Change (INECC)</li></ul>
Ecology and Climate Change	Compliance and Safe Environment Rights Improvement in Carbon Emission Reduction and Greenhouse Effect Components Rules for Mitigation and Adaptation to Climate Change Reduction and Vulnerability of Ecosystems	<ul style="list-style-type: none"><li>• National Institute of Ecology and Climate Change (INECC)</li></ul>

### *Evaluations Compliance, Quality and Budget*

According to the Federal Budget and Treasure Act a performance evaluation is carried out by verifying objectives and goals level of compliance based on strategic and management objectives, which allows to know the results of federal resources (Cámara de Diputados del H. Congreso de la Unión, 2006, p. 66). Additionally, this act establishes that the ministries and public organisations in charge of evaluating performance evaluation will perform evaluations by external evaluators which comply with ministries or organisations' requirements, objectivity and disclosure whatsoever stated by the applicable law.

Being an improvement management accountability tool, the evaluation requires robust methodology, compliance and highly measured quality standards, not only to give public administration results but also to achieve effective and efficient results, which may contribute to a democratic quality (AEC, 2010).

Concerning evaluations quality in the social, educational and environmental sectors, legal frameworks and organisations—CONEVAL, INEE and INECC—regulate evaluation and issue public policy improvement recommendations; coordinate activities among federal, state and local authorities in public functions and use strategic and management indicators to measure objectives achievement of social programmes; disseminate information (social sector); ensure quality outsourcing services to recommend techniques on evaluation instruments and their application and results (educational sector); and issue government policies at all levels (environmental sector).

### *Evaluation Use (Sectoral Level)*

As mentioned above, the evaluation seeks to support public policy decisions and inform citizens by allowing, on the one hand, timely correction and improvement of public interventions and alignment of budget design and control, considering the quality and impact of government actions, as well as validating and generating consensus on policies that require greater continuity with the presidential cycles (Ferreiro & Silva, 2010). Thus, evaluation is considered to be an articulating and cross-cutting mechanism capable of linking the formulation of public policies with their implementation, managing allocated resources and social impact effects

(CEPAL n.d.), hence the most common uses given to evaluation in the three sectors of national development already mentioned are:

1. Suggestions and recommendations **to improve public policies.**
2. **Decision-making on** public policy programmes, actions and/or strategies.
3. **Accountability** on government officials' responsibility actions.

Despite this, in Mexico there is not enough literature related to the use of evaluation information. One of the first efforts in this area was made by Meza (2017), who measured the instrumental use of evaluation in Mexico based on three indicators: (a) the percentage of recommendations issued by external evaluators that were selected as Aspects Susceptible for Improvement (ASM) by federal public officials; (b) the percentage of ASMs that were concluded; and (c) the percentage of ASMs that are considered adequate to improve the performance of federal Pp.

Notwithstanding these uses, much remains to be done to ensure that evaluation is at the forefront of public policy strategies and decisions in all areas of national development.

### *Integration of Evaluation into Parliamentary Structures*

According to the Fiscal Coordination Act, the SHCP with the technical support of the CONEVAL, annually submits a report to the House of Representatives on the adjustments made to the performance indicators and, if applicable, their justification, as well as a report on the progress achieved by the federal states, state and local demarcations with respect to the implementation and operation of the results-based budget (PbR) and Performance Evaluation System as it relates to the exercise of federal resources, in order that the Congress carries out a global evaluation.

Likewise, the SHCP submits a report every three months to the House of Representatives about the goals achieved by the programmes.

### *Annual Evaluation Process in Government Agencies*

The General Guidelines for the Evaluation of Federal Programmes of the Federal Public Administration establish that the CONEVAL, the SHCP and the Ministry of Public Function (SFP) are enforced to publish the

Annual Evaluation Programme (PAE) in which the types of evaluations that will be carried out to different budgetary programmes are defined, as well as their scheduling. On the one hand, the Evaluation Units (EU) of all federal agencies and entities are responsible for notifying the operational areas of the programmes that will be evaluated in each case, as well as hiring the external evaluators who will carry out the evaluation processes. These evaluators are generally academic and research institutions, legal persons or bodies specialised in this field (CONEVAL & SHCP, 2019 ).

The evaluations carried out to Public Policies can be in terms of design, consistency and results, processes, impact, strategic, complementary or specific.

Once the external evaluations have been concluded, the results are presented to the areas responsible for the programmes, as well as to the EU agencies and organisations so that, by mutual agreement, the Aspects Susceptible for Improvement are selected, which must be implemented in order to optimise, or where appropriate, correct the results of the evaluated programmes. Likewise, the results of the evaluations are published on agencies' and organisations' websites.

On the other hand, the Federal Budget and Treasury Accountability Act establishes that the CONEVAL and the SHCP are in charge of issuing an Annual Programme of Evaluations, where types of external evaluations are defined and will be carried out on transferred states and local funds based on strategic and management indicators, as well as schedules to perform evaluations (Cámara de Diputados del H. Congreso de la Unión, 2006).

In the social sector, CPEUM establishes that the CONEVAL is responsible for evaluating social development policy programs, objectives, goals and actions, as well as issuing recommendations and establishing liaisons with the federal, state and local authorities required to perform its duties. Until recently, the INEE was responsible for providing educational authorities in the educational field—federal, state and local, as well as the private sector—with the tools to carry out the evaluation of elements that comprise their educational systems. This institute will be replaced by the Teaching Revaluation National Institute for and Continuous Improvement in Education (Instituto Nacional para la Revaloración del Magisterio y la Mejora Continua de la Educación). Additionally, in terms of ecology and climate change, the INECC will be responsible for evaluating and

issuing recommendations on climate change, environmental protection and preservation and ecological balance restoration.

According to Meza (2017), approximately 1,700 external evaluations of federal budgetary programmes were carried out from 2007 to 2013. Of these, about 98% were design, consistency and outcome evaluations, performance specific and others, and only 2% were process and impact evaluations. According to this author, in this period 68% of the evaluations were specific-performance evaluations, which show progress in the objectives and goals fulfilment of the programmes through the analysis of results indicators, services and management on the summary of information provided by the agencies responsible for the programmes. The possible reasons for this type of evaluation being conducted mostly in this period are the level of maturity of the national evaluation policy in Mexico; insufficient information from Pp—mainly from the base—the cost and time involved in conducting process and impact evaluations; and policy changes.

### *Relationship Between Internal and External Evaluations*

The evaluation can be, according to the evaluator, internal or external. In the first one, the evaluator has a close relationship with the agency or organisation and therefore finds less resistance in obtaining information. However, internal assessments can be also considered as biased as they are carried out by the staff of the units or public entities. In the regard of the external evaluation “it contributes to increase evaluation compliance and transparency due to the objectivity assumed by the external team” (Gonzalez and Ozuna, as cited in Cardozo & Mundo, 2012, p. 33).

In Mexico, for example, the Federal Budget and Tax Accountability Act establishes that the SHCP must carry out the performance evaluation of the programmes based on the General Guidelines for the Evaluation of Federal Programmes of the Federal Public Administration. The SHCP and the CONEVAL publish the Annual Evaluation Programme of the Federal Programmes and the Federal Contribution Funds for each Fiscal Year, allowing for a more complete vision of governmental action (Cámara de Diputados del H. Congreso de la Unión, 2006).

According to this Law, the performance evaluation could be internal or external through specialised persons with proven experience in the subject matter to be assessed, who meet the requirements of independence, impartiality, transparency and others set out in the applicable provisions.

### *Subnational Monitoring and Evaluation Systems*

As mentioned earlier, Mexico is a federal republic, composed of 32 federal entities. Each federative entity is composed of municipalities. In Mexico there are more than 2,500 municipalities. Since the panorama of Mexico would not be complete if the advances achieved in the subnational area were not mentioned.

As of 2008, federal entities and municipalities were incorporated into Results Management, Results-Based Budget and Performance Evaluation System (PbR-SED) through adjustments to the federal legal framework that corresponds to them. From 2010 to 2018,<sup>1</sup> the SHCP has published each year a Diagnosis on the progress of the implementation of the Management for Results and the Performance Evaluation System of the states and municipalities (SHCP UED, 2019). Interesting results have emerged from these annual reports, the main one being that it has become a motivating element for subnational governments to carry out actions aimed at raising their rating in that evaluation of their progress in the implementation of the PbR-SED, as it has been observed frequently when carrying out field work in the federative entities. The assessment of the degree of progress is determined by means of a survey that is addressed to the 32 federal entities and 64 municipalities (selected based on their population), which contains quantitative assessments for the responses, with which an index is constructed that goes from zero to 100.

In the 2018 document of “Diagnosis of progress in the implementation of the PbR-SED in the states and municipalities” (SHCP UED, 2019), states like Mexico, Guanajuato and Baja California topped the list in 2018 with scores of 96.9 to 98.3 (in a scale of 100), demonstrating the interest of these states to strengthen their PbR-SED capabilities in recent years. On the other side are entities such as Guerrero (40.2) and Baja California Sur (33.0).<sup>2</sup>

<sup>1</sup> In 2019 the methodology was changed and only an analysis of the strengths, weaknesses, opportunities and threats (SWOT) of each federative entity and municipality was published on a provisional basis. For 2020, it is now in development an update of the methodology previous used.

<sup>2</sup> Document regarding compliance with provisions contained in the third paragraph of the Article 80 of the General Accounting Law Governmental. Secretariat of Finance and Public Credit. 2018. Also known as “Diagnosis of progress in the implementation of the PbR-SED in the states and municipalities”.

The topics that are evaluated, by the SHCP in the mentioned Diagnosis, to know the degree of progress in the implementation of the PbR-SED in the states and municipalities are the following (SHCP UED, 2019):

1. PbR-SED, which is composed of:
  - a. Legal framework. It is analysed the harmonisation of the municipality/state regulatory framework to federal regulations regarding PbR and performance evaluation system.
  - b. Planning. It is analysed the efforts made by subnational governments to include in their State Development Plans (PED) a results-oriented structure, with programmes that have clear indicators and goals that also allow monitoring their performance.
  - c. Programming. It is analysed that the public budget be linked to public programmes with indicators and challenging goals. This implies the definition of programmatic structures.
  - d. Budgeting. It is analysed that governments have public spending efficiency policies, as well as established criteria for the preparation and homogeneous presentation of financial information, accounting records and financial information on the income of resources and exercise of expenditure.
  - e. Exercise and control Analyses that the normative and organisational elements exist and are applied to guarantee the correct execution and efficiency in the exercise of the expense.
  - f. Monitoring. It is analysed that there are and apply regulatory and organisational elements that allow monitoring of the budget year.
  - g. Evaluation. It is analysed the existence of mechanisms, regulations and instruments necessary for the application of Evaluations that allow obtaining objective information on budgetary policies and programmes.
2. Transparency. It is analysed that there is and applies a set of rules that regulate that citizens can access information on the different stages of the PbR-SED, as well as the application of public resources.
3. Capacities Development. It is analysed that the states have training activities in the area of PbR-SED for the best implementation in the state/municipality.



4. Acquisitions. It is analysed that the state or municipality have established processes for the purchase of goods and the contracting of public services.
5. Human Resources. It is analysed that the state or municipality has elements that allow professionalisation and guarantee the permanence of human capital in the public service.

Of these sections those that, at the state level, have been better evaluated with (1) Acquisitions, (2) Capacities Development and (3) PbR-SED. Within the PbR-SED section, the category with the highest rating in 2018 was Legal Framework, Budgeting and Exercise and control.

At the municipal level, they have been better evaluated with (1) Acquisitions, (2) PbR-SED and (3) Transparency. Within the PbR-SED section, the category with the highest qualification in 2018 was Budgeting, Legal Framework and Planning.

## DISSEMINATION OF EVALUATIONS TO SOCIETY

### *About the Civil Society Organisations*

The non-profit sector, as mentioned above, is a heterogeneous set in political, social and ideological terms. Girardo and Mochi identified three characteristics shared by Civil Society Organisations (CSOs) (Girardo & Mochi, 2012).

- The first of these is an institutional dimension defined as the *non-state public*. This means CSOs are subjects acting on behalf of a common good, acting as a public strategy that opens and builds the public agenda.
- A second characteristic is that they seek to serve the community by focusing the attention on non-monetary values, which the state is not always able to provide based on public policies. The work of CSOs has placed these issues on the public agenda.
- Another characteristic of CSOs is the resources available they share. Donations and volunteer human capital.

There are not enough studies and/or statistical records that allow us to know more about CSOs, which prevents us from obtaining information on the issue at hand. Information can be known about CSOs that stand

out in awards, that call for public events to publish the results of their studies or to debate on interest topics, which are recognised in the media and social networks. However, a minority compared to the more than 40.000 existing CSOs has been registered in (National Institute of Social Development, INDESOL) as a donating organisation authorised by the SHCP.

### *Institutionalised Use of Evaluations by Civil Society*

Discussion about the civil society in Mexico refers to a wide range of multi-sector actors on different areas. This melting pot ranges from small associations that seek to support a local and perfectly defined cause, such as those large associations that professionally focus their efforts on addressing social problems parallel with national public policies. The evolution that the social sector had in Mexico to reach the current setting, in which there is a legal and fiscal framework, dates from the last 30 years (Cortés, 2010).

More recently, in 2004, the Federal Law for the Promotion of the Activities of CSOs was published, which established the creation of a Federal Registry of CSOs that could be subject to support for the promotion (Ministry of Social Development, 2018).<sup>3</sup> By the end of 2018, there were 42.269 registered CSOs in Mexico (National Institute of Social Development, 2019).

In Mexico there are civil society organisations and *think tanks* that use the evaluation and results of third-party evaluations as an essential element for the analysis of government actions.

An example of a common use of evaluations to public programmes is *Gestión Social y Conocimiento* (Social Management and Knowledge, GESOC A.C.) which is a non-profit analysis and research centre, specialised in generating methodological solutions and public evidence, open and robust, for the maximisation of the public value produced by the public and private sectors (GESOC, 2019a). Since 2009 (GESOC, 2019b) they publish the Performance Index of Federal Public Subsidy Programmes annually (GESOC, 2019c), whose objective is to evaluate

<sup>3</sup> Articles 2, 3.2 and 3.3.

the performance of federal social programmes, for this index the information of evaluations is used and made every year in the context of Performance Evaluation System by external evaluators and CONEVAL.

Another outstanding example is México Evalúa, a centre of knowledge and analysis that focuses on the evaluation and monitoring of government operations to raise the quality of their results (México Evalúa, 2019). The centre has published public spending, anti-corruption, security and justice and education documents. It conducts evaluations, reviews evidence generated by other bodies and makes recommendations.

Annually, the CONEVAL awards recognitions to “good practices in the use of monitoring and evaluation results in the public policy cycle” (CONEVAL, 2019a). In this way, it seeks to promote and disseminate the best social development practices implemented based on monitoring actions or results found in evaluations.

There are four awards categories:

- Outstanding actions for generating evidence
- Outstanding practices in the use of information and evidence
- Highlighted exercises to improve monitoring activities
- Outstanding trajectories in the generation, analysis and use of evidence

These awards are given to federal and sub-national government agencies, civil society organisations, and even to the media that participate in the process of generating, analysing and using evidence that contributes to decision-making, as well as promotion of evaluation (CONEVAL, 2019b).

In 2018, the CSO that obtained recognition for having performed outstanding actions for the generation of evidence was: Acción Ciudadana Frente a la Pobreza, for its drive to generate and document evidence for social development (CONEVAL, 2019b).

In the category of outstanding practices in the use of information and evidence, the work of Plan Estratégico de Juárez, A.C. was recognised for the design and implementation of the project ‘Así estamos Juárez’; and CREA Comunidades de Emprendedores Sociales, for the use of evidence to implement changes for improving the programme Mujeres Moviendo a México (CONEVAL, 2019c).

The Espinoza Yglesias Study Centre received an award for an outstanding track record in the generation, analysis and use of evidence to generate, transmit and disseminate evidence for social mobility (CONEVAL, 2019c).

In other years, other CSOs have been recognised, such as: México Evalúa (2015), Acción Mexicana contra la Pobreza (2015), ¿Cómo Vamos? (2015), GESOC, A.C. (2015), Centro de Estudios Espinoza Yglesias (2015) (CONEVAL, 2019a).

Although there are CSOs that have already internalised the use of evaluations as standard practice, and these results are visible, for the vast majority there is no documented information.

### *Public Opinion and Discussion of Evaluation Findings*

In Mexico, the culture of evaluation is proliferating among different segments of the population. Since 2015, the Evaluation Week, promoted by public and academic institutions, has been held in Mexico.<sup>4</sup> While the target audience for these events has traditionally been people from the public sector, academics and civil society, it has increasingly permeated more sectors of the population. Subnational governments have been interested in organising events to publish results of actions undertaken in evaluating public programmes and policies.

The media (TV and specialised journals) are also an important channels for dissemination and discussion of evaluation findings. There are discussion TV and radio programmes<sup>5</sup> that invite academics, civil society and officials to discuss on issues of public interest for considering evaluations results.

In the annual awards ceremony held by CONEVAL, media have also been recognised. A television station and its reporters; newspapers and a radio group obtained an award for the use of the information generated by CONEVAL. These media reach more sectors of the population making known the results of evaluations depending on the context.

<sup>4</sup> The Evaluation Week is a meeting point in which the public sector, civil society and the academic community participate with the organisation of different activities to contribute to research and discussion on the importance of monitoring and evaluation for the continuous improvement of public policies and programmes. In 2015 and 2016 a national event was carried out in México and from 2017 it has been held in Latin American and the Caribbean (CLEAR LAC n.d.).

<sup>5</sup> For example, Primer plano, México Social, Dinero y Poder, Línea directa, Agenda Pública, among others.

On the other hand, the Budget Transparency web site (SHCP UED, 2019) contains evaluations results performed as part of the Performance Evaluation System of 2007 and 2018, which are also downloadable. Recently, CONEVAL has promoted on different social networks and other audio-visual media, information resulting from evaluations into a language that is accessible to the public. In addition, CONEVAL holds conferences, seminars and round tables, both face-to-face and transmitted over the Internet, in which different actors of the evaluation ecosystem are invited to discuss.

### *Role of Society*

Currently, Mexico is going through a transition due different changes in the Federal Public Administration. Federal Government actions have opened a space for public debate and deliberation. Most recent public consultations for decision-making have raised reactions from different points of view, there are followers of the current President who defend and support his decisions without any question, and others who demand evidence-based decision-making.

The society that demands decision-making based on evaluations<sup>6</sup> reflects a greater degree of knowledge about the existence of evaluations and their usefulness as evidence to support decision-making. This has not yet permeated into the society. A lot of work must continue to be done in terms of evaluation and creation of instruments that carry, in a simple language, the main findings and conclusions on the design and results of social and non-social programmes.

## PROFESSIONALISATION

### *Academic Offer*

In the Mexican Educational System, the upper level of studies comprises 3 levels: Higher University Technician, Bachelor and Postgraduate. The postgraduate course contains three degrees that are Specialty, Master's and Doctorate (Ministry of Education, 2019).

<sup>6</sup> So far, even before the official inauguration, there have been two consultations with the public. The first on the location of the New Mexico City Airport; and the second on the Mayan Train construction and the implementation of ten social programmes.

The specialty degree is aimed at the training of individuals trained for the study and treatment of specific problems of a particular area of a profession, being able to refer to knowledge and skills of a basic discipline or to specific activities of a particular profession, and has as background Academic bachelor's degree (Ministry of Education, 2019).

The Master's degree is aimed at the training of individuals trained to participate in the analysis, adaptation and incorporation into practice of the advances of a specific area of a profession or discipline, and have at least as academic background the bachelor's degree or specialty (Ministry of Education, 2019).

The Doctorate degree is aimed at the training of individuals trained for teaching and research, with mastery of particular subjects of an area. Graduates are able to generate new knowledge independently, or to apply the knowledge in an original and innovative way, and have at least the academic background of the bachelor's degree, specialty or master's degree (Ministry of Education, 2019). In Mexico, evaluators have originally been trained in various disciplines and academic degrees, such as public administration, political science, economics, sociology, anthropology, among others. It is through postgraduate courses that they acquire the knowledge and skills necessary to strengthen the practice of evaluation.

There are also master's and postgraduate programmes that offer in their curricula a significant content of subjects, in their professorships, oriented towards the training of evaluators, such as: qualitative methods, quantitative methods, statistics, econometrics, impact evaluation, public policy analysis, among others.

The major specialty, master's, and doctoral programmes are shown in Table 10.3.

Several institutions offer postgraduate programmes under diplomas and courses modalities, some of them are mentioned in Table 10.4.

Furthermore, other institutions, mainly consultancies, offer courses and workshops on topics such as evaluation, indicators construction, logical framework model and, occasionally, quantitative methods.

**Table 10.3** Postgraduate programmes academic offer (own development)

<i>Academic Institution</i>	<i>Programme</i>	<i>Modality</i>
Public Health National Institute	Specialisation in Comprehensive Evaluation of Social Development Policies and Programmes	Online
Latin American Faculty of Social Sciences	Specialisation in Educational Policy and Management	Full time
Centre of Economic Research and Teaching (CIDE)	Master's in administration and Public Policy	Full time
Latin American Faculty of Social Sciences	Master's in Social Sciences	Full time
Centre of Economic Research and Teaching (CIDE)	Master's in Economics	Full time
Autonomous Technological Institute of Mexico (ITAM)	Master's in Applied Economics	Full and part time
Latin American Faculty of Social Sciences	Master's in Energy and Environmental Management	Full time
National Autonomous University of Mexico (UNAM)	Master's in Government and Public Affairs	Full time
Latin American Faculty of Social Sciences (Flasco)	Master's in Government and Public Affairs	Full time
Centre of Economic Research and Teaching, A.C. (CIDE)	Master's in Methods for the Analysis of Public Policies	Online
Latin American Faculty of Social Sciences (Flasco)	Master's in Population and Development	Full time
Latin American Faculty of Social Sciences (Flasco)	Master's in Social Development Policy and Management	Full time
Latin American Faculty of Social Sciences (Flasco)	Master's in Compared Public Policy	Remotely
Latin American Faculty of Social Sciences (Flasco)	Master's in Public Policy Compared Public Policy and Gender	Remotely
Latin American Faculty of Social Sciences (Flasco)	Ph.D. in Social Sciences Research	Full time
National Autonomous University of Mexico (UNAM)	Ph.D. in Political and Social Sciences	Full time
Centre of Economic Research and Teaching, A.C. (CIDE)	Ph.D. in Public Policy	Full time

While not all programmes are clearly named to reflect that they are an evaluators programme, a review of the curricula discloses all of them are oriented towards evaluation or to strengthen some aspects of the evaluators' profile.

**Table 10.4** Diplomas and courses offer (own development)

<i>Academic Institution</i>	<i>Programme</i>	<i>Modality</i>
ACEVAL	Diploma in Public Programmes Evaluation and Design	Online *Only for ACEVAL associates
Centre CLEAR LAC	Impact Evaluation Course with Statistics	Part time
Centre CLEAR LAC	Processes Evaluation Course	Part time
Centre CLEAR LAC	Qualitative Tools for Programmes Evaluation Course	Part time
Centre CLEAR LAC	Quantitative Tools for Programmes Evaluation Course	Part time
Centre CLEAR LAC	Intro to Programmes Evaluation Course	Part time
Centre CLEAR LAC	Diploma in Qualitative Evaluation of Public Policy	Part time
Centre CLEAR LAC	Diploma in Impact Evaluation	Part time
Centre CLEAR LAC	Diploma in Public Policy and Evaluation	Part time
Centre of Economic Research and Teaching, A.C. (CIDE)—National Laboratory of Public Policy	School of Methods: different courses	Part time
Centre CLEAR LAC	Evaluation Workshops Impact Evaluation Programmes of Population, Health and Nutrition	Full time
Centre CLEAR LAC	Evaluation Workshops Qualitative Methods for Evaluating Health Programmes	Full time
Mexico College	Summer workshops	Full time
National Autonomous University of Mexico (UNAM)	College of Mexico	Online

### *Evaluation as a Profession*

#### *Association of Evaluators*

In Mexico in 2014, the Mexican Evaluators Society (ACEVAL) was formalised as a non-profit civil association. It is equivalent to what is known internationally as VOPE (Volunteer Organisations for Professional Evaluation). It started formal activities in October 2014 with ten founding members.

ACEVAL has the following four purposes (ACEVAL, 2019a):



- Bring together evaluators, officials involved in the evaluation, academics and, in general, population interested in evaluation, and to be an exchange of experiences and knowledge on the subject.
- Promote the professionalisation of the people involved in public offices that demand evaluations and people who offer their services as evaluators.
- Create an observatory of evaluations of projects, programmes and public policies in Mexico.
- Be a national benchmark in evaluation.

In 2015, International Year of Evaluation, ACEVAL's objective was to become known nationally and internationally, in order to grow its membership and position itself. That year, the Evaluation Week was held for the first time in Mexico and was the basis to launch ACEVAL in Mexico. At an international level on that year, representatives of ACEVAL participated in events such as the IV ReLAC Conference (Peru), the Global Evaluation Forum (Nepal), REDLACME-COPLAC Conference (Panama), and in training events at the Universities of Carleton (Canada) and Antwerp (Belgium). ACEVAL grew its membership by 60% that year. The main topic in ACEVAL's round table was the *Market of Evaluators*.

In 2016, ACEVAL partnered with government and academic institutions as well as held events open to the evaluation community, grew its membership more than 100% over the previous year. In 2017 it contested to host the 5th ReLAC Conference and it was selected. It received a recognition from the University of Antwerp for best performance in the strengthening implementation strategy designed during 2015. This year, the main topics in ACEVAL's round table were *the Evaluators' Market, the use of evaluations* and *the evaluators' profile*.

In 2017, ACEVAL hosted the Joint Evaluation Conference organised by ReLAC—REDLACME—IDEAS at the University of Guanajuato, which meant a strong coordination work with hosting networks and co-hosting national institutions. This year ACEVAL had a greater positioning in the evaluation community, which was reflected in an increase in membership to 85 members. This year, the central topics in ACEVAL's discussion table were the *Evaluators Market, the use of evaluations, other evaluation methodologies, evaluation at subnational levels*, and it started working with the international EvalYouth initiative.

In 2018, ACEVAL (ACEVAL, [2019a](#)) took a turn in its working scheme. Seven lines of work were defined, three national and four coming

from international initiatives, specifically from EvalPartners (EvalPartners, 2019):

- **Evaluation market** National line of work in 2016 a diploma was designed and driven with a solid component in professionalisation. In 2017, for the first time, it was given to ACEVAL members and in 2018 the second edition of the diploma was shared with the Salvadoran Evaluation Network. Collaboration agreements were signed with state government bodies to support, strengthen and help them promote the evaluation culture. Scholarships were sought for members to take courses and diplomas in academic institutions.
- **Use of evaluation** This line of work seeks to disseminate the evaluation culture to promote the use of evaluation results. To this end, representatives of ACEVAL are invited to participate in seminars or conferences of other institutions to give lectures on this subject.
- **Evaluation methodologies** A new team on socio-economic assessment of investment projects was built within this line of work. This new group held events during Evaluation Week and has established links with government areas in charge of this issue.
- **EvalYouth** This line seeks the development of new generations of evaluators. In 2018, a critical reflection and redesign of work schemes was carried to reorient them and in 2019, programmes aimed at achieving this objective will be implemented through links with academic institutions and consulting firms in evaluation for the development of a mentoring programme.
- **EvalGender** work team was created this year, and events of its own and others have been held in coordination with academic bodies and other CSOs, with the aim of disseminating the incorporation of a gender perspective in public programmes and their evaluations.
- **EvalParlamentarios** This work team was created this year with the aim of involving legislators in evaluation issues, contributing to their knowledge on the subject and promoting the use of evaluation results in decision-making. ACEVAL participated in EvalColombo and International Forum of EvalParlamentarios.
- **EvalSDG** This year also saw the creation of a task force to promote the evaluation of ODS. It is the youngest working group and is still developing its work programme.

Currently, ACEVAL has more than 90 active members (ACEVAL, 2019b) and three honorary members. This has allowed the creation and functioning of working groups and numerous face-to-face events in the capital city and headquarters inside the country, as well as virtual events. ACEVAL is already a national benchmark in evaluation, both by the evaluation community and by federal and sub-national government agencies, as well as academic institutions.

### *Evaluators Profile*

In Mexico, there are consulting firms focused on evaluation, academic and research institutions that participate in evaluations, and independent evaluators. The Budget Transparency (SHCP UED, 2019) web site publishes the evaluations carried out in accordance with the PAE, which show evaluation plurality among bodies that carry them out.

There is no certification as evaluators in the country, and there is no consensus among the main actors on the need to generate certified evaluators. Institutions such as CONEVAL has developed a solid criterion to define evaluators technical profile to be hired.

The legal framework in evaluation includes the Ministry of Finance and Public Credit and the CONEVAL, both of which coordinate evaluations contained in the PAE, and oversee reviewing evaluations.

### *Compliance and Quality Obligations*

Evaluators should adhere to the evaluation's terms of reference, which normally establish information confidentiality and criteria under which evaluations should be conducted.

There is no code of ethics for evaluators, however ACEVAL is currently working in one for its members, and as soon as it is published it will be sent to the regulatory bodies to disseminate and adhere to it.

ACEVAL, as part of the Monitoring Network, Evaluation and Systematisation of Latin America and the Caribbean (ReLAC), has adopted the Evaluation Standards published by ReLAC. However, it has not been greater dissemination about these standards.

In the Mexican M&E system there is no arbitration figure for the evaluations. The SHCP and CONEVAL are responsible for reviewing the evaluations. There is a mechanism called 'Institutional Positioning' that allows the units evaluated to argue nonconformities, and discard any conclusions or recommendations and that these are not incorporated into the "Susceptible Aspects of Improvement" (CONEVAL and SHCP, 2019).

## CONCLUSION

In the last decade, the evaluation in Mexico had a generalised impulse thanks to the creation of a legal framework that allowed the institutionalisation of the evaluation of public resources in the orders of federal, state and local government.

This legal framework, together with the various information systems to monitor the results of the public programmes, laid the foundations for promoting a culture of evaluation with increasingly strong methodologies for measuring government performance.

It should be noted that the sectors that showed the greatest progress in the area of evaluation were social development, education and ecology and the environment, thanks to the design of legal frameworks and the creation of governmental bodies responsible for coordinating and regulating the evaluation in these areas, such as they are the National Evaluation Council (CONEVAL) responsible for measuring poverty and evaluating programmes created for the promotion of the development of human and social capital; the creation of the INEE that focused, in the educational field, its efforts in the creation of a National System of Educational Evaluation in Mexico (SNEE) through which the National Evaluation Policy was oriented of Education (PNEE); or the INECC that allowed progress in the evaluation of Mexico's climate policy, consolidating the evaluation as a feedback and learning tool for the improvement of public policies at the national level.

This culture of evaluation included government agencies, in the three orders of government, federal, state and municipal, which are required to publish, year after year, an Annual Evaluation Programme of their different Pp and convene, for such purposes, to academic and research institutions, legal persons or specialised agencies in the field of evaluation for the development of evaluations of different types: design, consistency and results, specific performance, processes, impacts, among others.

For its part, the institutionalised use of evaluations by various CSOs and reflection groups has contributed to the impulse of an evaluative culture since they use evaluation and its results as a fundamental means to analyse and monitor to government actions.

Another advance in the field of evaluation is the multiple post-graduate programmes (specialty, master's and doctorate) offered in the country by various academic institutions of great recognition, such as the Center for Economic Research and Teaching Center (CIDE),

the Faculty Latin American Social Sciences (FLACSO Mexico) or the National Autonomous National University of Mexico (UNAM), among many others, that offer academic programmes, almost entirely, face-to-face and full-time, in addition to the specialisation programmes offered by organisations of civil society, academic institutions and consulting. Likewise, the link between professionals, academics, officials and the general population dedicated to evaluation thanks to the creation of spaces for reflection, debate, training and training promoted by organisations such as the National Academy of Evaluators of Mexico (ACEVAL), constitute another important achievement for the improvement of the evaluation.

However, and despite these advances, in Mexico there is not enough literature to know the use of the information derived from the evaluation, nor is there a consensus on the need to generate certified evaluators that allow us to continue advancing, in a determined way, in the generalisation of the evaluation and the use of its results.

Regarding sub-national levels, since the end of the last decade, M&E systems began to be developed in several states that are now pioneers and have evolved with robust information systems to monitor their indicators and reach their goals.

Finally, it is important to mention that in Mexico a great challenge is currently being faced for the SED. In 2018 there was a change in the federal government of Mexico, the new president disapproves of the SED model implemented by the previous government; however, this system continues because it is in the regulatory framework. Internally and due to the institutionalisation, that exists and despite the change in the headlines of the regulatory units (Ministry of Finance and CONEVAL), there have been no substantial changes in the Evaluation System. There are nine priority programmes of the new government, which are being evaluated by CONEVAL in their design and implementation, this work is being done in coordination with the executing agencies. This shows institutional strength beyond changes in national policy.

### LIST OF ABBREVIATIONS

ACEVAL	Mexican Evaluators Society
ASM	(System) of Aspects Susceptible for Improvement
CONEVAL	National Council for Evaluation of Social Development Policy
CSO	Civil Society Organisations
INDESOL	National Institute of Social Development
INECC	National Institute of Ecology and Climate Change

INEE	National Institute for the Evaluation of the Education
LGDS	General Law of Social Development
M&E	Monitoring & Evaluation
PAE	Annual Evaluation Programme
PbR	Results-Based Budget
PbR-SED	Results-Based Budget and Performance Evaluation System
Pp	Budgetary Programmes
SHCP	Ministry of Finance and Public Credit

## REFERENCES

- ACEVAL. (2019a, December). *Annual reports*. <http://aceval.org/wp-content/uploads/2014/12/Informe-ACEVAL-2014-2017.pdf>. Accessed December 10, 2019.
- ACEVAL. (2019b, December). *Us*. <http://aceval.org/nosotros/>. Accessed December 10, 2019.
- AEC. (2010). *Fundamentos de evaluación de políticas públicas*. Ministerio de Política Territorial y Administración Pública.
- Cámara de Diputados del H. Congreso de la Unión. (2004). *Ley General de Desarrollo Social*. Cámara de Diputados del H. Congreso de la Unión. Sitio web: [http://www.diputados.gob.mx/LeyesBiblio/pdf/264\\_250618.pdf](http://www.diputados.gob.mx/LeyesBiblio/pdf/264_250618.pdf)
- Cámara de Diputados del H. Congreso de la Unión. (2006). *Ley Federal de Presupuesto y Responsabilidad Hacendaria*. Cámara de Diputados del H. Congreso de la Unión. Sitio web: [http://www.diputados.gob.mx/LeyesBiblio/pdf/LFPRH\\_200521.pdf](http://www.diputados.gob.mx/LeyesBiblio/pdf/LFPRH_200521.pdf)
- Cardozo, M. (2006). *La evaluación de políticas y programas públicos. El caso de los programas de desarrollo social en México*. Cámara de Diputados, LIX Legislatura.
- Cardozo, M., & Mundo, A. (2012). *Guía de orientación para la evaluación de políticas y programas de desarrollo social*. Consejo de evaluación del desarrollo social del df.
- CEPAL (n.d.). *Acerca de Evaluación de políticas y programas públicos*. <https://www.cepal.org/es/temas/evaluacion-de-politicas-y-programas/acerca-evaluacion-politicas-programas-publicos>. Accessed April 23, 2019.
- CLEAR LAC. (n.d.). *Evaluation week 2015*. <https://www.clear-la.org/semana-de-la-evaluacion/>. Accessed December 8, 2019.
- CONEVAL, & SHCP. (2019, December). *General guidelines for the evaluation of federal programs federal public administration*. <https://www.coneval.org.mx/Evaluacion/NME/Paginas/LineamientosGenerales.aspx>. Accessed May 22, 2019.

- CONEVAL. (2019a, December). *Best practices 2015*. [https://www.coneval.org.mx/Evaluacion/BPME/GF/Paginas/BP\\_2015.aspx](https://www.coneval.org.mx/Evaluacion/BPME/GF/Paginas/BP_2015.aspx). Accessed May 22, 2019.
- CONEVAL. (2019b, December). *Best practices 2018*. [https://www.coneval.org.mx/Evaluacion/BPME/Documents/Resultados\\_BP\\_2018.pdf](https://www.coneval.org.mx/Evaluacion/BPME/Documents/Resultados_BP_2018.pdf). Accessed on 22 May 2019.
- CONEVAL. (2019c). *Buenas Prácticas en el Uso de los Resultados de Monitoreo y Evaluación en el Ciclo de las Políticas Públicas 2019*. CONEVAL. [https://www.coneval.org.mx/Evaluacion/BPME/Documents/Resultados\\_BP\\_2019.pdf](https://www.coneval.org.mx/Evaluacion/BPME/Documents/Resultados_BP_2019.pdf). Accessed May 22, 2019.
- Cortés, L. E. (2010). *Una fotografía de la Sociedad Civil en México*. CEMEFI. Diario Oficial de la Federación. (1997). Decreto por el que se crea la Coordinación Nacional del Programa de Educación, Salud y Alimentación, como órgano desconcentrado de la Secretaría de Desarrollo Social. *Secretaría de Gobernación*. [http://www.dof.gob.mx/nota\\_detalle.php?codigo=4890287&fecha=08/08/1997](http://www.dof.gob.mx/nota_detalle.php?codigo=4890287&fecha=08/08/1997). Accessed 2019.
- EvalPartners. (2019, December). *About us*. <https://www.evalpartners.org/about/about-us>. Accessed December 10, 2019.
- Evaluare. (2019, January). *¿Para qué se evalúan los proyectos, programas y políticas públicas?* <https://www.evaluare.mx/para-que-se-evaluan-los-proyectos-programas-y-politicas-publicas/>. Accessed January 12, 2019.
- Ferreiro, A., & Silva, F. (2010). Evaluación del impacto y calidad de las políticas públicas: Hacia una agencia independiente. *Libertad y Desarrollo: Serie Informe Económico*, 203, 3–30.
- GESOC. (2019a, April). *GESOC, A.C.* <http://www.gesoc.org.mx/page-16/>. Accessed April 17, 2019.
- GESOC. (2019b, April). *History*. <http://www.gesoc.org.mx/page-16/page-e-23/>. Accessed April 17, 2019.
- GESOC. (2019c, April). *INDEP*. <https://www.indep.gesoc.org.mx/>. Accessed April 17, 2019.
- Girardo, C., & Mochi, P. (2012). Las organizaciones de la sociedad civil en México: Modalidades del trabajo y el empleo en la prestación de servicios de proximidad y/o relacionales. *Economía, Sociedad y Territorio*, 12(39), 333–357.
- Kliksberg, B., & Rivera, M. (2007). *El capital social movilizado contra la pobreza: la experiencia de comunidades especiales en Puerto Rico*. CLACSO.
- Larrú, J. M. (2010). Algunas cuestiones conceptuales y metodológicas en torno a la evaluación de impacto. *Revista E-Valuación*, 11, 20–31.
- México Evalúa. (2019, December). *Homepage*. <https://www.mexicoevalua.org>. Accessed December 10, 2019.

- Meza, J. (2017). *Principales factores asociados al uso instrumental de la evaluación externa de los programas presupuestarios federales en México, periodo 2007–2013*. Doctoral dissertation, Universidad nacional autónoma de México.
- Ministry of Education. (2019, December). *The structure of the Mexican education system*. [https://www.sep.gob.mx/work/models/sep1/Resource/1447/1/images/sistemaedumex09\\_01.pdf](https://www.sep.gob.mx/work/models/sep1/Resource/1447/1/images/sistemaedumex09_01.pdf). Accessed December 10, 2019.
- Ministry of Social Development. (2018). Operation rules for the social co-investment program. *Official Journal of the Federation* (DOF).
- National Institute of Social Development. (2019). *Directory of civil society organisations registered in the federal register of CSOs 2019*.
- SHCP UED. (2019, April). *Transparencia presupuestaria*. <https://www.transparenciapresupuestaria.gob.mx/>. Accessed April 17, 2019.





## Evaluation in Peru

*Brenda Bucheli del Águila, Susana Guevara Salas,  
and Emma Lucía Rotondo Dall’Orso*

### GENERAL OVERVIEW

Peru has been a democratic, social, independent and sovereign republic, continuously since 1980. Its government is unitary, representative and decentralised and it is organised in three main branches: executive, legislative and judiciary, following the principle of separation of powers. Its political democratic system is constituted as a State and it is ruled by the Political Constitution approved in 1993. It is a presidential regime and it counts also on a Parliament, which is elected every four years.

In this context, evaluation in Peru has mainly been developed within the last two decades, following Latin American trends also aiming at this goal, as pointed out by Chianca and Younker (2004). These authors have

---

Founding partners of the Peruvian Evaluation Network EvalPerú.

---

B. Bucheli del Águila (✉) · S. Guevara Salas · E. L. Rotondo Dall’Orso  
EvalPerú, Lima, Peru

B. Bucheli del Águila · S. Guevara Salas  
Pontificia Universidad Católica del Perú, Lima, Peru

highlighted an evident rise regarding the use of professional evaluation among several sectors (public, private and non-profit), in the beginning of the twenty-first century, as well as the emergence of national and regional organisations of evaluators, specialised publications and short-run and postgraduate education programmes.

The trends identified at that initial moment are totally applicable to the Peruvian case, as shown by some important milestones, such as: (1) the emergence of the Peruvian Evaluation Network (EvalPerú) in 2002, (2) the organisation of the first ReLAC (Monitoring, Evaluation and Systematisation Network for Latin America and the Caribbean) conference in our country in 2004, or (3) the launching of the first diploma course on Monitoring and Evaluation of Social Projects and Programmes in the Pontifical Catholic University of Peru (PUCP) (EvalPerú, 2018c; PUCP, 2017).

Evaluation is born in Peru from social sciences, including Economics, Sociology and Anthropology, as well as Psychology, Education and even Medicine, different sources from which it takes approaches and methodologies, which are recreated for its applied purposes. Its practice is linked to the trends guiding the design and management of development policies, programmes and projects (PPP) and it searches for greater efficiency and transparency in development interventions. In the last 20 years, while Peru has been leaving the ‘developing country’ status and becoming a middle-income country trying to join the Organisation for Economic Co-operation and Development (OECD), foreign aid available resources for PPP have also progressively decreased. They are being replaced by public resources (at least partially), which are becoming increasingly important. As a consequence, the non-governmental organisation (NGO) sector is decreasing, as government interventions become more relevant for the country’s development.

The aforementioned situation allows a better understanding of the increasing involvement and interest professionals across the public sector show towards evaluation, as well as the incorporation of a new legal framework required for its implementation in public institutions. The reorganisation of state entities, which is a key element regarding the inclusion of evaluation within their organisational structure and the creation of specialised areas in this field becomes clear.

In Peru, the practice of evaluation within public administration is recent. It began in 2000, within the framework of government reform and modernisation. The implementation of Budgeting for results (PpR)

in 2007, promoted by the Ministry of Economic and Financial affairs, constituted a milestone marking the development of evaluation in the country. This forced all the other sectors of the three government levels (national, regional and local) to implement PpR as well. In this context, monitoring and evaluation development became important. Specifically, evaluations increased in the last years.

The creation of the Ministry of Development and Social Inclusion (MIDIS), in late 2011, represented a second milestone. Its structure included monitoring and evaluation of social programmes and projects. This sector shows big progress, not only regarding implemented evaluations, but also, their results.

Meanwhile, corruption has been gaining ground in the Latin-American region and in the country, which has demanded further efforts to fight against it. The situation caused by Odebrecht's mismanagement illustrates this (La República, 2017). Even though the quest for more transparency regarding the use of resources and the fight against corruption are important for the institutionalisation of evaluation in the country and among civil society, this potential has not been exploited, as it will be shown.

## STRUCTURES AND INSTITUTIONAL PROCESSES (POLITICAL SYSTEM)

### *Evaluation Regulations*

In general, legal norms establishing the need and mandatory character of evaluation in all government levels and sectors (except for the parliament<sup>1</sup>) have been enacted in the country, as described below:

- The Framework law for the modernisation of the public administration (Law N. 27658, enacted on March 29th, 2002) highlights evaluation as a key action in the modernisation process of the public administration. Its institutionalisation assures control over actions implemented by the State.

<sup>1</sup> There is no centralised monitoring and evaluation system for the Executive Branch institutions in Peru. It rather emerged from the results-based management proposal made by the Ministry of Economic and Financial affairs and then it spread into other sectors, especially the social sector. The Legislative branch has not integrated any monitoring and evaluation process.

- The Organic law of regional governments (Law N. 27867 enacted on November 18th, 2002) regards evaluation as one of the guiding principles of regional policies and management in the framework of modern management and accountability.
- The General National Budget System Law (Law N. 28411 approved on December 6th, 2004) establishes that public expenditure policies must be orientated towards the fulfillment of macro-fiscal stability and that they have to be executed throughout results oriented management, with efficiency, efficacy, economy and quality. Likewise, it states that budget modifications must “include sufficient and appropriate information for the evaluation and monitoring of goals and targets” (article VIII). The norm includes budget evaluation regarding results and financial performance as well, which “represents a source of information for the budget planning stage, matching public expenditure quality improvement” (article 46).
- The Executive branch organic law and amendments thereto (Law N. 29158, enacted on December 19th, 2007), establish that one of the duties of the Executive branch is to evaluate national and sectoral policies. Likewise, it considers that national and sectoral policy evaluation is one of the competencies of the Ministries, in coordination with regional and local governments.
- The Public administration Budget Law for 2007 (Law N. 28297, enacted on December 12th, 2006) establishes a results oriented budget methodology, which was progressively implemented in the whole national, regional and local state apparatus. The goal of this norm was to restructure budget processes in order to incorporate information concerning performance, resource tracking, mandate, obligations and results evaluation throughout indicators.
- National policies of compulsory fulfilment for National government entities (Supreme decree N. 027-2007-PCM approved on March 22nd, 2007. Norm repealed by Supreme decree N. 032-2018-PCM on March 23rd, 2018). The norm established that Ministries should annually approve the performance goals and targets of a six-monthly evaluation of policy fulfilment.
- Creation of the Strategic Planning National System and the Strategic Planning National Centre (Legislative decree N. 1088 enacted on June 28th, 2008). Its duties include the monitoring and evaluation of plans, policies, goals, programmes and priority national development projects.

- The regulation of National Policies (Supreme Decree N. 029-2018-PCM, enacted on 19th March, 2018) establishes the competence of monitoring and evaluation of the government's general policy and its sectoral policies as well. The Presidency of the Council of Ministers and the Ministries are held responsible for evaluation, respectively.

Considering the national normative provisions, different sectors have created instances in charge of monitoring and evaluation. In addition, they have developed norms which guide their evaluation actions.

The first attempts to implement monitoring and evaluation systems in the Ministry of Economic and Financial Affairs (MEF) date from 1995. They were supported by the Investment Office, once it was granted the planning duties, transferred from the dismantled National Planning Institute in 1992. Between 1995 and 2000, the Investment Office was in charge of analysing investment proposals presented by different sectors. It enacted the Investment Law and, from 2001 on, all public investment projects had to comply with certain norms regarding the project cycle in order to be approved. From 2006, an Investment Monitoring and Tracking Operative System (SOSEM) was implemented (Shack, 2007).<sup>2</sup>

In parallel, between 1996 and 1998, the Financial Management Integrated System (SIAF) was created in order to control the execution of public expenditure. Since 2000, provisions regarding the development of Multi-year Sectoral Strategic Plans, Multi-year Public Investment Programmes and Institutional Strategic Plans, including their monitoring and evaluation components, were established (Shack, 2007).

Since 2004, a Public Expenditure Monitoring and Evaluation System was developed, in order to “provide timely, reliable and relevant information, which allows to decide on the allocation and execution of scarce resources with a view to maximise social well-being” (Shack, 2007, p. 15). This system came into operation the following year and it was progressively implemented within the three government levels.

Since 2007, the aforementioned Public Sector Budget Law integrated evaluation instruments into the public budget. The Public Budget National Direction was now in charge of monitoring and evaluation.

<sup>2</sup> In the first years, some ministries used the terms ‘monitoring’ and ‘evaluation’ to design different processes. Since 2010, a tendency to use the term ‘tracking’ is discernible.

The implementation of the PpR methodology, since 2007, represents a milestone regarding the institutionalisation of monitoring and evaluation systems, because it “introduced a change concerning budgeting, based upon an integrated vision of planning and budgeting and the articulation of actors and actions oriented towards the achievement of results. PpR focuses on the resolution of critical issues affecting the population. It represents the crossover from a budgeting process focusing on the execution of means to a budgeting process characterised by the achievement of ends” (Shack & Rivera, 2017, p. 16).

Talledo (2015) adds that “the implementation of PpR has significantly evolved (in the last years) and it has opened to the integration of budget programmes. Nowadays, according to the memorandum contained in the Public Sector Budget Law regarding fiscal year 2013 (Law 29 951), more than 48% of non-financial and social security public expenditure is oriented towards results (throughout a programmatic wording based upon a Logframe). Likewise, several independent evaluations regarding diverse budgetary programmes have been carried out” (p. 320).

The Ministry has a board in charge of independent evaluations within the Public Budget National System, under the PpR umbrella (Board N. 009-2008-EF/76.01 amended by Board Resolution N. 023-2012-EF/50.01). This norm defines evaluation and two types of evaluation are established (Evaluation of Budget Design and Execution–EDEP and Impact Evaluation–EI). Likewise, guidelines regarding evaluation development, actors involved, reports, recommendations and commitment, entered into by the instance in charge of the budget to be evaluated and the MEF and its monitoring, are provided.

The MIDIS is the governing institution regarding social development, inclusion and equity, and it is responsible for policies, programmes and social projects on the three government levels.

The law by which the MIDIS was created sets monitoring and evaluation of the performance and goals attained by social development policies, plans and programmes at the national, regional and local levels (Law 29792 enacted on October 20th, 2011) among its competences. This fact represented a milestone in evaluation history, since, for the first time, a state entity explicitly announced it.

Later, in 2012, this entity approved the guidelines for evaluation, monitoring and management of evidence regarding policies, plans,

programmes and projects carried out by the MIDIS (Ministerial Resolution N. 192-2012-MIDIS enacted on October 23rd, 2012). This document establishes the concepts, standards, general and specific procedures regarding the development of monitoring, evaluation and management of evidence of policies, plans, programmes and projects within the sector.

It also contains specific provisions on monitoring and evaluation, and minimum standards for the evaluation of policies, plans, programmes and projects carried out by the MIDIS; as well as provisions on the development of the Annual Evaluation Plan and its types (design evaluation, process evaluation, results evaluation, impact evaluation, systematic reviews, meta-analysis of impact evaluations, ex ante evaluations and specific evaluations), the development of methodological notes on evaluations, oversight and publication of evaluations.

The Ministry of Women and Vulnerable Populations (MIMP) is the guiding institution of the sector oriented towards women, children and adolescents, older adults, people with disabilities and internal migrants.

In 2004, the norms and procedures of the National Monitoring and Evaluation System (SIME) concerning women and social development (Ministerial Resolution N. 331-2004-MIMDES) were approved. Later, in 2011, the norms and guidelines for sectoral monitoring and evaluation were approved (Ministerial Resolution 412-2011-MIMDES). Both documents were oriented towards the monitoring and evaluation of social problems constituting the ground in response to which the MIDIS was created, in 2012.

More recently, the MIMP approved the Monitoring and Evaluation Norms (RM N. 142-2016-MIMP enacted on 27th June, 2016) and derogated previous norms. The current norms aim at establishing roles, responsibilities, procedures and tools for monitoring and evaluation of policies, plans, programmes and projects carried out by the MIMP.

The aforesaid document points out specific provisions for monitoring and evaluation. As for evaluation, provisions set out the conditions to be fulfilled in order to carry out an evaluation, namely: an ‘evaluability’ analysis, a conceptual note, terms of reference and an evaluation matrix. Evaluation standards are also established.

### *Evaluation Practice*

Evaluations carried out in the public sector have been encouraged by Ministries as guiding entities of each sector. Each sector develops annually

an evaluation plan, which is approved together with the public budget for the next year.

Within the framework of Law N. 28927 regarding the Public Sector Budget for fiscal year 2007, the MEF started to implement two types of evaluations: EDEP and EI. EDEP are evaluations which focus mainly on the design of the budgetary programme and analyse implementation, efficiency, efficacy and quality of goods and services provided to the population to a lesser extent. Impact Evaluations analyse the effects of interventions and the degree to which these effects can be explained as a result of the aforementioned interventions, based upon an experimental methodology. In both cases, evaluations are external, since they are carried out by independent evaluators at the expense of the Public Budget General Direction. This instance counts the development of monitoring and evaluation of public expenditures (Regulation of Organisation and Duties, approved by Supreme Decree N. 117-2014-EF and amended by Supreme Decree N. 221-2016-EF on July 22nd, 2016) among its duties.

Independent evaluations have been carried out since 2008, after the aforementioned PpR process was established. To this end, guidelines orienting independent evaluations of the public budget are developed. Between 2008 and 2018, 58 EDEP were carried out in 16 sectors, among which education was the one concentrating most evaluations, followed by health, justice, social inclusion and development and agriculture (see Table 11.1).

In 2012, the MEF started to implement impact evaluations. Between that year and 2018, the MEF promoted the development of 15 impact evaluations, among which in 2018 eight were already finished and seven are still in progress (see Table 11.2).

Between 2012 and 2018, the MIDIS, created in 2011, carried out two types of evaluations: impact evaluations (12) and design, implementation or results evaluations (20 in total). All evaluations have concerned social programmes and social policy instruments deployed by the Ministry. Evaluations are external, carried out by researchers not linked to the institution.

This Ministry has a social innovation lab called AYNi Social Lab. It intends to “identify and implement innovative solutions in order to improve the quality of life of populations living in poverty or vulnerability conditions, according to the social policy priorities established by the government. These solutions will be adopted by social programmes



**Table 11.1** Ministry of Economic and Financial affairs. Number of evaluations of budget design and execution (EDEP) by sector/per year (Ministerio de Economía y Finanzas, 2018a)

<i>Sector</i>	<i>2008</i>	<i>2009</i>	<i>2011</i>	<i>2013</i>	<i>2015</i>	<i>2016</i>	<i>2017</i>	<i>2018</i>	<i>Total</i>
Agriculture			1	1	1	1		1	5
Environment				1			1		2
Foreign trade and Tourism			1						1
Culture			1						1
Development and Social inclusion	2				1	1	1		5
Economic and Financial affairs			1	1					2
Education	1	3		5	1			1	11
Energy and Mining		1							1
Justice, Internal Affairs, Prosecutor		1	2	3					6
Women and vulnerable populations			1		1				2
Presidency of the Council of Ministers		1					1		2
Production				1					1
Health		1	1	2	4				8
Labour and employment promotion		1				1			2
Transportation and Communications	1	1	1					1	4
Housing, Construction and Sanitation		1	2	2					5
Total	4	10	11	16	8	3	3	3	58

in order to be furtherly implemented, evaluated and to analyse their scalability” (Ministerio de Desarrollo e Inclusión Social, 2018a).

Social innovations are evaluated following experimental methods, which allows to estimate the causal effect directly attributable to the tested solution. In 2018, there were five evaluations in progress (Table 11.3).

The instance in charge of evaluation within the MIDIS is the Monitoring and Evaluation General Direction, a line agency of the institution, whose duties are established in the Regulation of Organisation and

**Table 11.2** Ministry of Economic and Financial affairs. Number of impact evaluations (EI) by sector and status (Ministerio de Economía y Finanzas, 2018b)

<i>Sector</i>	<i>Finished</i>	<i>In progress</i>	<i>Total</i>
Agriculture		1	1
Development and Social Inclusion	2	2	4
Economic and Financial Affairs		1	1
Education	2	3	5
Justice, Internal Affairs, Prosecutor	1		1
Justice, Internal Affairs, Prosecutor			
Health	2		2
Labour and employment promotion	1		1
Labour and employment promotion			
Total	8	7	15

**Table 11.3** Ministry of Social Inclusion and Development. Number of evaluations carried out by type and per year (Ministerio de Desarrollo e Inclusión Social, 2018b)

<i>Evaluation type</i>	<i>2012</i>	<i>2013</i>	<i>2014</i>	<i>2015</i>	<i>2016</i>	<i>2017</i>	<i>2018</i>	<i>Total</i>
Impact evaluation	1	2		2	2	5		12
Design/performance/results evaluation			3	2	8	6	1	20
Total	1	2	3	4	10	11	1	32

Duties (approved by Supreme Decree N. 006-2017-MIDIS, article 54). Among other duties, it is in charge of establishing guidelines, methodologies, standards and technical criteria for evaluation and conducting the design and oversight of design, implementation and impact results evaluations regarding national and sectoral policies, its instruments and social programmes concerning social inclusion and development.

According to available information, the MIMP encouraged the development of nine evaluations between 2015 and 2017, regarding national multi-sectoral plans (gender, children and adolescents, family, older adults and people with disabilities) and projects implemented in the sector as well (Table 11.4). All of them were external evaluations.

In 2012, the MIMP created the Policy Monitoring and Evaluation General Office, as a counseling instance subordinated to the General Secretary of the Ministry, in charge of evaluation. The regulation of

**Table 11.4** Ministry of women and vulnerable populations. Number of evaluations carried out by type and per year (Ministerio de la Mujer y las Poblaciones Vulnerables, policy monitoring and evaluation office<sup>a</sup>)

<i>Evaluation type</i>	<i>2015</i>	<i>2016</i>	<i>2017</i>	<i>Total</i>
Evaluation of a national multi-sectoral plan	1	2	1	4
Project evaluation		3	2	5
Total	1	5	3	9

<sup>a</sup> Answer to an inquiry addressed by e-mail to Susana Guevara in July, 2018.

Organisation and Duties assigned it the duty of “proposing norms and guidelines for the implementation and effective and efficient functioning of the integral monitoring and evaluation of national and sectoral policies” (ROF approved by Supreme Decree N. 003-2012-MIMP, article 34).

Since 2007, the evaluation of the educational system is the main concern of the Ministry of Education (MINEDU). Therefore, the Quality Learning Measurement Office (UMC) designed and implemented national evaluations on the learning achievements of basic education students. Evaluations are census-based and international evaluations of sampling character are implemented. Evaluations are carried out annually and they involve second-grade students from public and private educational institutions. In some provinces, evaluations target fourth-grade students having a mother tongue other than Spanish, attending Bilingual Intercultural Schools. Since 2016, learning achievements made by second high-school year students are evaluated (Ministerio de Educación, 2016).

The Strategic Monitoring and Evaluation Office, which makes part of the Strategic Planning Secretary, in the MINEDU, is “responsible for coordinating the process of production, integration and analysis of statistic data concerning the performance and impact of educative policies” (MINEDU, 2015, p. 39). In 2016, this office created the cost-effective innovation lab Minedu LAB, “a tool for innovation and learning, which allows the design, implementation and evaluation of improvements addressed to the policies following experimental methods, in order to determine their effectiveness before scaling, at a very low cost for the sector” (Ministerio de Educación, 2018). Within this framework,

six experimental evaluations have been implemented, three are still in progress and five are being designed.

The **Ministry of Agriculture and Irrigation** has a Policy Monitoring and Evaluation General Direction, a line agency subordinated to the ministerial office in charge of sectoral evaluations. Among its duties, it has to “conduct the impact evaluation of national and sectoral policies, sectoral plans, programmes, special projects and norms within the three government levels” (Regulation of Organisation and Duties, approved by Supreme Decree N. 001-2017-MINAGRI, article 49).

Meanwhile, in 2017, the **Ministry of Production** created the Impact Evaluation and Economic Studies General Office, a counseling body. Among its duties, it has to conduct and promote the development of studies and baseline measurements, intermediate and impact evaluations regarding sectoral interventions, as well as the design of corresponding methodologies (Regulation of Organisation and Duties, approved by Supreme Decree N. 002-2017-PRODUCE, article 41).

In 2014, the **Ministry of Housing** created the Monitoring and Impact Evaluation General Office, a counseling body within the Ministry. Among its duties, it has to develop norms and guidelines for the implementation and effective and efficient operation of monitoring and integral impact evaluation of national and sectoral policies within its competence. It also has to design and establish monitoring and impact evaluation strategies addressing projects and programmes, as well as the management and implementation of national and sectoral policies within the sector.

### *Use of Evaluations*

As mentioned before, the MEF orients the development of evaluations in order to improve the quality of public expenditures. Thus, once EDEP were finished, the evaluated body and the MEF agree on performance improvements throughout a commitment matrix based upon the recommendations of the evaluation. Bodies report the compliance with commitments to the MEF, annually, and it forwards a report on the implementation of improvements to the Congress of the Republic (Ministerio de Economía y Finanzas, 2015).

Out of all EDEP carried out between 2008 and 2017, eleven matrices had still to be signed and 14 evaluated public interventions did not lead to signed matrices since they were discontinued interventions (Ministerio de Economía y Finanzas, 2017). 422 commitments regarding 17 sectors

and 409 means of verification developed out of all signed matrices. Out of this total, “513 means of verification have been fulfilled, which represents a 72.9% level of progress. As mentioned before, means of verification lie at the core of the evaluation of commitment compliance reached by evaluated institutions. In this respect, all means of verification regarding a commitment must be fully fulfilled in order for it to be considered complied with” (Ministerio de Economía y Finanzas, 2017, p. 11).

Aiming at promoting the use of results delivered by the monitoring and evaluations system, the MIDIS developed an ‘evidence management’ approach, which “timely supports the arguments of decision makers, based upon accurate data” (Velásquez, 2016, p. 85). The steps to be completed in order to implement evaluation management include elaboration of recommendations resulting from the evaluation, analysis of their validity, application of viable recommendations and improvement proposals.

The ‘Guidelines for the Evaluation, Monitoring and Management of Evidence’ concerning policies, plans, programmes and projects implemented by the MIDIS specify the patterns to be followed in order to develop the recommendations arising from evaluations. Once the evaluation is finished, the Monitoring and Evaluation General Direction analyses the viability of recommendations, as well as their politic, normative, institutional, technical, economic and social implications, together with managers and technicians representing the body whose programmes were evaluated. An agreement and commitment proposal is elaborated for the implementation and monitoring of said recommendations.

Available data show that the MIDIS established two commitment matrices regarding its programmes in 2017, and one in 2018. Their level of progress has not been made public.

The MIMP establishes patterns for the implementation of recommendations arising from evaluations in the ‘MIMP Monitoring and Evaluation Norms’. This institution establishes Commitment Matrices to be signed by the Policy Monitoring and Evaluation Office and the evaluated body, no later than 20 days after evaluation is finished. Commitment matrices established in the last year are not available to the public.

## DIFFUSION IN SOCIETY AND ACCEPTANCE (SOCIAL SYSTEM)

### *Institutionalised Use of Evaluations Made by Civil Society*

In Peru, evaluation is not considered a public good allowing to inform society about social action carried out within the framework of public policies. It is still being developed within the organisational and administrative field, at best, in order to provide feedback for public administration decisions. There is no law, norm or regulation stating the necessity of disseminating or discussing evaluations among civil society, since the current approach is mostly administrative.

The mechanism used to disseminate evaluations is based upon institutional websites or digital repositories belonging to the bodies promoting evaluation, whether public or foreign aid entities. It is also unknown whether these channels include all developed evaluations or just some of them, or if evaluation reports are fully or partially issued.

It is important to specify that the concept of evaluation use is restricted, both in the public and private sector, to programme management, reaching external aid sources and, exceptionally, programme or project stakeholders, but not the general public.

In this way, the chance evaluation offers to reflect and discover what is working or not in order to reach the desired impact is underestimated. Neither is it considered as an opportunity for social learning regarding a public intervention. The use civil society makes of evaluation is minimum. The main underlying obstacle is the idea that evaluation concerns private and administrative use, rather than being a public good, since this idea prevails both in the public and private spheres. The attempts to develop evaluation capacities in the public sphere have been scarce, generally delimited to the training of evaluation professionals (PREVAL 2000–2012 and USAID 2013–2018) and to the improvement of the internal capacities of technical groups making part of projects financed by foreign aid, in order to design and develop monitoring and evaluation systems.

### *Perception and Public Debate on Evaluation and Its Results*

Evaluation is considered neither as a tool to improve public policy nor as a means to evaluate its effectiveness and efficiency among civil society

institutions concerned with social dialogue. Even though public institutions (such as the MEF, MIDIS and MIMP) voluntarily post evaluations on their websites or digital repositories, the way in which these are used by the general public remains unknown.

The first obstacle is the pre-conception of evaluation as a trans discipline subject to programme management and not as a public good regarding social learning. This perception prevails among experts and users, maybe because this topic is not addressed in evaluation training.

### *Demands for Evaluation Arising from Civil Society*

In Peru, there are two spaces for institutionalised social dialogue: the Roundtable for the Fight Against Poverty (MCLCP) and the National Agreement.

The MCLCP was created in 1981, constituted as a concertation space between civil society and the state, in order to improve national, regional and local public policies. Its action is integrated into the main contents of the '*Carta Social*' (Social Charter) and the goals and duties contained in the decrees giving rise to it (MCLCP, 2001).

The MCLCP is characterised by the combination of both organisation and territorial representation criteria, as well as sectoral-institutional criteria. This is shown in the organisation chart and the structure of its different levels. At a territorial scale, the MCLCP counts on a national body, 26 regional instances and several local (provincial and district) instances as well. There are also regional roundtables, which do not match political delimitations and articulate diverse territories: interprovincial, inter districts, basin-related, composed in this way because of the geographical and socioeconomic reality of territories.

Roundtables represent different civil society and state instances and they aim at institutionalising the participation of the general public in the design, decision-making and audit process regarding public social policy; maximising transparency and integrity in management and reaching higher efficiency in the implementation of programmes aiming at the fight against poverty (MCLCP, 2018). Even though citizen participation is promoted as an inherent characteristic of the discussion on policies and public affairs, so that they count on oversight committees on social programmes, in general, no roundtable includes evaluation, neither as an agenda topic, nor as an input for debate.

Regarding the National Agreement, it is a social dialogue forum where the three government levels (national, regional and local) participate, together with other important political and social institutions in the country. It began in 2002, with the signature of a summary document on public policies, which were classified into the following categories: strengthening of democracy and rule of law, equitable development and social justice, promotion of the country's competitiveness and affirmation of an efficient, transparent and decentralised state (Acuerdo Nacional, 2018). The use of evaluation is mentioned, neither in an agreement, nor in mission or vision statements.

Regarding civil society, there are two kinds of organisations that could demand the debate and the use of evaluations. On one hand, NGOs grouped into the National Association of Research and Social Promotion Centres (ANC), created in the 1980s, gathering more than a hundred NGOs and thematic observatories. ANC is a guild body gathering Peruvian NGOs. Among its objectives, there is "the impact on proposals addressing integral development and improvement of the quality of democracy in the country, from a rights, gender equality and interculturality perspective" (ANC, 2018).

Nevertheless, ANC has existed for already 30 years, its articulation of proposals aiming to impact on public policy does not demand neither the use of evaluations, nor spaces for the debate, analysis and reflection on public policy evaluation. This subject is addressed neither in its institutional mission statement, nor in the projects it implements searching to impact on policies.

For instance, ANC's project called 'Strengthening of the National Association of Centres in order to promote the broadest participation of civil society in the country's democratisation and development', which is financed by the European Union, does not even quote evaluation as an input for debate. This project aims at strengthening the capacities of the platform and ANC networks, in partnership with the Foreign Aid International Entities Coordination Agency (COEECI), in its role of guilds and development agents articulating proposals, building partnerships and impacting on public policy in Peru in favour of human development. Among the results, there is the consolidation of national dialogue spaces, where ANC networks and COEECI participate together with the state. None of its objectives and actions quote evaluation, since its mobilising potential regarding transparency in public administration is not appreciated (ANC, 2017).



COEECI is a network founded in 1994, gathering 52 private aid organisations. They do not use evaluation in its public emissions, documents and observatories, and they do not even demand a debate on it. Its main goal is to act as an organised speaker representing these entities before the Peruvian state and before related private and public institutions as well.

COEECI seeks to become an exchange and reflection space on subjects interesting development actors, and to reach coordination levels in order to contribute to national development efforts as well. Even though its explicit mission is “to be a speaker before Private Foreign Aid, Civil Society and the Peruvian state, as well as to reach coordination levels in order to contribute to national development efforts”, its objectives do not mention evaluation. For example, its sixth objective is “to guide or commission research and develop working documents on development in Peru, which can serve as a basis for the exchange of dialogue, ideas and proposal development” (COEECI, 2018).

In short, there is no institutionalised debate among the State and Civil Society counting on evaluation as an input. Evaluations are voluntarily placed upon Foreign Aid demands or because public entities post them on their websites, the use made out of these evaluations remains unknown.

## PROFESSIONALISATION

### *Academic Courses and Training*

The beginnings of specialised education date from the early twentieth century, where specific training and the development of support material regarding monitoring and evaluation were offered by public and private development programmes and projects, in order to achieve more efficient interventions. In this initial context, the Programme for the Strengthening of Monitoring and Evaluation Regional Capacities regarding Reduction of Rural Poverty in Latin America and the Caribbean International Fund for Agricultural Development (IFAD) projects (PREVAL) is an outstanding actor, because it assumed the development of local evaluation capacities as its mandate and it specially benefited Peru and Colombia.

During its period of validity, between 1995 and 2008, PREVAL contributed to the on-service training of the executors of projects financed by IFAD, as well as evaluation practitioners representing Civil

Society and the Government. It accompanied several non-academic training and short duration activities, such as workshops and exchanges of experiences, developing additional basic material in Spanish, such as a glossary and a guide (Chianca & Youker, 2004; FIDA, 2002, 2006).

The current offer of higher education programmes oriented towards the training of evaluators in Peru comprises specialised postgraduate programmes, offered by Peruvian and international academic institutions in face-to-face and online learning mode. Six of them have been identified.

The first programme of this kind emerged in 2004. The specialised area of Social work within the Faculty of Literature and Human Sciences, in PUCP, launched that year the Diploma Course on Monitoring and Evaluation of Social Projects and Programmes, answering to the request made by the graduates of the Diploma Course on Design and Management of Social Projects, also offered at that moment twelve face-to-face and six on-line cohorts have finished the course ever since; 400 professionals, mainly belonging to the public sector (in the last years) and counting on previous experience in monitoring and evaluation have been trained (PUCP, 2017, 2018).

Years later, between 2013 and 2017, a new diploma course came into being and was developed by the Peruvian University Cayetano Heredia (UPCH), supported by the US Agency for International Development USAID within the framework of the '*Evaluations*' Project. The entry point for the design of the programme was constituted both by the diagnosis implemented by this project among USAID's partner organisations on individual (among professionals fulfilling monitoring and evaluation duties) and institutional capacities (based upon organisational performance in the planning, monitoring, evaluation and knowledge management areas), (UPCH, 2018; USAID, 2017). Since 2018, UPCH offers the aforementioned diploma course in blended learning mode, open to any person interested.

Also, in this year, a new diploma specialisation course on monitoring and evaluation was offered at ESAN University (ESAN, 2018). This special programme was offered on July 25th 2018 (Table 11.5).

Furthermore, in recent years, technological innovations have facilitated the emergence of online international programmes in Peru, such as three Masters Degrees recognised by the Universities offering them, but not by the government of the country hosting them (Masters '*Propios*' in

**Table 11.5** Monitoring and/or evaluation postgraduate programmes offered to evaluators in Peru between 2016 and 2018 (own development)

<i>Nº</i>	<i>Name</i>	<i>Type</i>	<i>Learning mode</i>	<i>Promotors</i>	<i>Country</i>
1	Monitoring and Evaluation of Social Programmes and Projects	Diploma course	Face-to-face/ on-line	PUCP	Peru
2	Monitoring and Evaluation of Development Plans, Programmes and Projects	Diploma course	Face-to-face/ on-line with a tutor	UPCH	Peru
3	Monitoring and Evaluation of Programmes and Projects	Diploma course	Face-to-face	ESAN	Peru
4	Design, Management and Evaluation of Public Policies and Projects	Master's Degree	Online	Universidad de Alicante/Asociación Argentina de Evaluación (Argentinean Evaluation Association)	Spain/Argentina
5	Evaluation of Public Policy	Master's Degree	Face-to-face/ on-line	Universidad Internacional de Andalucía	Spain
6	Evaluation of Public Policy	Master's Degree	Online	Universidad de Sevilla	Spain

Spanish) (which amount to diploma courses in Peru), offered by Spanish Universities: Alicante, Andalucía and Sevilla.

Moreover, Peruvian evaluators have a series of short duration training courses at their disposal, mainly specific and face-to-face, with different complexity levels and accents, developed by academic entities or other private entities, such as NGOs or Foreign aid institutions. By way of illustration, some examples are included in Table 11.6.

**Table 11.6** Monitoring and evaluation specialisation courses offered to Peruvian evaluators between 2016 and 2018 (own development)

<i>Nº</i>	<i>Name</i>	<i>Type</i>	<i>Learning mode</i>	<i>Promotors</i>	<i>Country</i>
1	Introduction to Impact Evaluation of Public Policies and Programmes	Course	Face-to-face/ On-line	CLAD	Panama
2	Evaluative Monitoring of Projects using a Results Oriented Management (EGR) and Log frame (EML) Approach through SIMER	Course	Online with a tutor	FAO	Chile
3	Evaluation of Development Programmes	Course	Face-to-face	GRADE	Peru
4	Monitoring and Evaluation of Social Programmes and Projects	Course	Online	REDINFOR	Peru
5	Impact evaluation of social programmes	Course	Online	Massachusetts Institute of Technology/JPAL	Chile
6	Using system concepts in evaluation design	Seminar—workshop	Face-to-face/ On-line	PUCP/EvalPeru	Peru
7	Project development and evaluation	Course	Face-to-face	PUCP/Centro Cultural PUCP	Peru

(continued)

**Table 11.6** (continued)

<i>Nº</i>	<i>Name</i>	<i>Type</i>	<i>Learning mode</i>	<i>Promotors</i>	<i>Country</i>
8	Impact Evaluation	Course	Face-to-face	Universidad del Pacífico	Peru
9	Introduction to Impact Evaluation	Course	Face-to-face	PUCP/CISEPA	Peru
10	Evaluation of Public Investment Projects within the framework of the SNIP	Course	Face-to-face	PUCP/Escuela de Gobierno y Políticas Públicas (School of Government and Public Policy)	Peru
11	Design and Evaluation of Public Policy	Course	Face-to-face	PUCP/Escuela de Gobierno y Políticas Públicas (School of Government and Public Policy)	Peru
12	Monitoring and Evaluation of Social Projects	Course	Face-to-face	PUCP/Escuela de Gobierno y Políticas Públicas (School of Government and Public Policy)	Peru
13	Impact Evaluation and Environmental Impact Studies within the Mining Strategic Sector	Course	Face-to-face	INTE-PUCP	Peru
14	Monitoring and Evaluation based upon Results-oriented management	Course	Online	UPCH	Peru
15	Public Policy Evaluation Expert	Specialisation course	Online	Universidad de Sevilla	Spain

Another source of information for Peruvian evaluators lies at educative programmes or courses focusing on the design and management of policies, programmes and projects, which usually include a module on indicators, monitoring and/or evaluation. The accent of these programmes or courses is wide and it might be oriented towards public investment or social projects related issues (Table 11.7).

**Table 11.7** Training programmes and courses containing modules on monitoring and/or evaluation offered to Peruvian evaluators between 2016 and 2018 (own development)

<i>Nº</i>	<i>Name</i>	<i>Type</i>	<i>Learning mode</i>	<i>Promotors</i>	<i>Country</i>
1	Performance indicators	Course	Face-to-face	PUCP/Escuela de Gobierno y Políticas Públicas (School of Government and Public Policy)	Peru
2	Design and formulation of social projects	Course	Face-to-face	PUCP/Escuela de Gobierno y Políticas Públicas (School of Government and Public Policy)	Peru
3	Management of development projects	Course	Online	IDB	No information available
4	Management of social projects	Diploma course	Face-to-face	UCSS/CIDIR	Peru
5	Management of social projects and foreign aid	Diploma course	Online with a tutor	Universidad Complutense	Spain
6	Public Investment Projects	Specialisation course	Face-to-face	Universidad del Pacífico	Peru

(continued)

**Table 11.7** (continued)

<i>Nº</i>	<i>Name</i>	<i>Type</i>	<i>Learning mode</i>	<i>Promotors</i>	<i>Country</i>
7	Formulation and Management of Public Investment and Local Development Projects	Specialisation course	Face-to-face	URP	Peru
8	Social management	Master's Degree	Blended	PUCP	Peru
9	Government and public policy	Master's Degree	Face-to-face	PUCP/Escuela de Gobierno y Políticas Públicas (School of Government and Public Policy)	Peru
10	Social management and human resources	Master's Degree	Face-to-face	UIGV	Peru
11	Government and public administration	Master's Degree	Face-to-face	UIGV	Peru
12	Social Policy/Management of Social Projects	Master's Degree	Face-to-face	UNMSM	Peru
13	Management of social programmes and projects	Master's Degree	Face-to-face	UPCH	Peru

In addition to what has already been mentioned, national evaluators complete their professional training by participating in national and international events periodically organised, for example, those organised by ReLAC. In 2015, the IV ReLAC conference took place in Lima, Peru, in coordination with PUCP and EvalPerú (ReLAC, 2015). In 2017, the X ReLAC Conference was jointly organised with RELACME and IDEAS in Guanajuato, Mexico. The Peruvian delegation counted 20 people, and this conference was lately spread in Lima (EvalPerú, 2018a).

Other examples of this non-academic training spaces are the 'Evaluation Week' of 2018, promoted by ReLAC, in which EvalPerú organised a panel discussion on evaluation standards for public policies (EvalPerú, 2018b), or the 'Evidence Week', both of which took place in 2016 and

2017 and were organised by a group of Peruvian partners interested in the subject, awakening interest beyond national borders (On Think Tanks, 2018).

### *Profession/Discipline*

There are two national networks in Peru, whose integration is voluntary, linked to the discipline: EvalPerú and the Peruvian Monitoring and Evaluation Network (REDPERUME). Both networks focus mostly on training activities, such as face-to-face events regarding exchange of experiences and discussion. EvalPerú organises online events, too, and it counts on an electronic interest list and a website.

The origins of EvalPerú date from the late 1990s, under the influence of PREVAL, which strongly contributed to the cohesion of ideas, people and interests around evaluation in our country and other countries in the region. A group of professionals engaged in evaluation started informal meetings in order to exchange experiences, contacts and new perspectives, and thus the idea of having a name, membership and making small contributions in order to fund outreach and training activities emerged.

In 2002, EvalPerú was born as a network of professionals interested in monitoring, evaluation and systematisation, without external funding and it has kept the same line ever since. It has been present in major moments of evaluation history, both nationally and internationally, such as:

- The foundation of ReLAC in 2003
- The first assembly of the global network of evaluation networks, International Organisation for Co-operation in Evaluation (IOCE), in 2004
- It has organised two regional ReLAC conferences in 2004 and in 2015
- It has contributed to the design of the monitoring and evaluation diploma courses in PUCP and UPCH.

Since 2013, it is formally established, counting currently on a management core of 13 people and more than 350 members, which are linked through different tools, such as an electronic interest list, a website and face-to-face and online events. It has signed agreements with institutions



such as USAID, IOCE and PUCP and it is currently affiliated to ReLAC and IOCE (EvalPerú, 2018c).

REDPERUME is a group which is attached to the Latin-American and Caribbean Monitoring and Evaluation Network (REDLACME), promoted by instances such as the World Bank (BM) and the Inter-American Development Bank (IDB). Its management team is constituted by the Monitoring and Evaluation Office of the MIMP, which mainly consists of public sector professionals working close to monitoring and evaluation in Peru.

The goal of this network is the strengthening of monitoring and evaluation culture in Peru, throughout spaces oriented to dialogue, discussion, outreach and development of capacities for all of its members. Since it was constituted in 2012, it has organised four specialised events. Currently, REDPERUME has 286 attached members and it has been supported by institutions such as USAID (REDPERUME, 2018). REDPERUME is one of the sponsors of the ‘Evidence Week’.

Furthermore, the country does not count on norms or principles regulating the professional exercise of evaluation, neither does it have independent arbitration mechanisms settling disputes specifically in this area, only regular channels before which all contractual disputes are brought.<sup>3</sup> Likewise, there is no formal professional certification system for evaluators, neither at individual nor at institutional level.

In the current situation, the forced professional relocation caused by the closure of several NGOs in the context of the changes the country has been going through, has led independent professionals with no training in monitoring and evaluation to offer their services in the area, individually or in association with consulting firms.

There are no exact figures on the concurrence among bidders in evaluation calls, but observation of people within this discipline shows that it is consistently high and chances arise mostly in the public sector, which is regulated by contracting regulations regarding this sector. Besides, in general, people demanding evaluations do not have the necessary and enough technical and solid background allowing them to identify the best

<sup>3</sup> Such as the norms contained in the Civil Code and the Judicial System, or the National Institute for the Defence of Competitiveness and the Protection of Intellectual Property.

proposals and low prices usually determine the final choice, especially in the case of meagre budgets destined to evaluations.<sup>4</sup>

There are no universally accepted quality standards regarding evaluation. However, it is worth highlighting that the OECD (Organisation for Economic Co-operation and Development) criteria are the most known and used when evaluations are requested, even though they are usually applied automatically and with no great analysis on their relevance.

As mentioned in the first section of this document, the form and contents of some evaluation products have been homogenised in several public instances, like the MEF, which represents an important step forward in the safeguard of evaluation quality.

In the academic field, postgraduate lecturers hold a master's degree level, since this is a condition set by the Higher Education National Superintendence. Nevertheless, the academic degree they hold belongs to other disciplines, since there is no evaluation postgraduate studies programme in the country.

There are no evaluation specialised journals in the country, either, as mentioned in Peruvian journals databases Scientific Electronic Library Online (ScIELO), Scopus, and Emerging Source Citation Index/Web of Science (ESCI/Wos) (ULADECH Católica, 2018).

### *Compliance with Quality Norms and Obligations*

As mentioned in the previous section, the country does not count on formally established universal quality norms or criteria, neither for evaluations, nor for evaluators. Those requesting evaluations usually include the guidelines for development and acceptance of the commissioned products in the terms of reference.

## CONCLUSION

As presented in this chapter, evaluation is relatively young in Peru, both as a discipline and as a practice. However, a notorious space among different sectors in society has been gained since the early twenty-first century, which has encouraged the development of a critical mass of professionals

<sup>4</sup> Reflections originating from internal discussions held among EvalPerú partners.

working in this field, as well as training programmes. Ever since, evaluation has arrived to stay, facing relevant changes in the national context. Its current validity is undeniable, especially regarding public and private social programmes.

Nevertheless, evaluation in Peru is still a work in progress, and important challenges remain, such as:

- Institutionalisation of evaluation: Evaluation shows different institutionalisation levels among branches, sectors and levels of government. Progress is observed regarding normativity, practice and use of evaluation in public institutions, especially those in the social sector, but there is only an emerging development of it within subnational and regional levels, and the subject is absent in the Congress of the Republic.
- Civil society has not assessed and assimilated evaluation as a public good for social learning, because evaluation is mostly oriented towards the private sector and management. Therefore, the chance to use evaluation in reflective and value-relevant processes regarding social learning on public interventions is lost.
- Even though some government instances count on guiding documents on the practice of evaluation, the development of norms or principles regarding the professional exercise of evaluation constitutes an important challenge.

In the face of these challenges, the development of an offer of higher education programmes oriented towards the training of evaluators within Peruvian universities, as well as the existence of national networks are opportunities to strengthen national evaluation capacities.

### LIST OF ABBREVIATIONS

ANC	Asociación Nacional de Centros de Investigación y Promoción Social (National Association of Research and Social Promotion Centres)
COEECI	Coordinadora de Entidades Extranjeras de Cooperación Internacional (Foreign Aid International Entities Coordination Agency)

EDEP	Evaluación de Diseño y Ejecución Presupuestal (Evaluation of Budget Design and Execution)
EI	Evaluación de Impacto (Impact Evaluation)
EvalPerú	Red Peruana de Evaluación (Peruvian Evaluation Network)
MCLCP	Mesa de Concertación para la Lucha contra la Pobreza (Roundtable for the Fight Against Poverty)
MEF	Ministerio de Economía y Finanzas (Ministry of Economic and Financial Affairs)
MIDIS	Ministerio de Desarrollo e Inclusión Social (Ministry of Development and Social Inclusion)
MIMP	Ministerio de la Mujer y Poblaciones Vulnerables (Ministry of Women and Vulnerable Populations)
MINEDU	Ministerio de Educación (Ministry of Education)
NGO	Non-Governmental Organisation
OECD	Organización para la Cooperación y el Desarrollo Económicos (Organisation for Economic Co-operation and Development—OECD)
PPP	Policies, Programmes and Projects
PpR	Presupuesto por Resultados (Budgeting for Results)
PREVAL	Programa para el Fortalecimiento de la Capacidad Regional de Seguimiento y Evaluación de los Proyectos FIDA para la Reducción de la Pobreza Rural en América Latina y El Caribe (Programme for the Strengthening of Monitoring and Evaluation Regional Capacities regarding Reduction of Rural Poverty in Latin America and the Caribbean FIDA projects)
PUCP	Pontificia Universidad Católica del Perú (Pontifical Catholic University of Peru)
ReLAC	Red de Seguimiento, Evaluación, y Sistematización de América Latina y El Caribe (Monitoring, Evaluation and Systematisation Network for Latin America and the Caribbean)
SIAF	Sistema Integrado de Administración Financiera (Financial Management Integrated System)
USAID	Agencia de los Estados Unidos para el Desarrollo Internacional (United States Agency for International Development)

## REFERENCES

- Acuerdo Nacional. (2018). *Acuerdo Nacional*. <https://acuertonacional.pe/>. Accessed November 1, 2018.
- ANC. (2017). *Proyecto CSO—LA/2016/377404-Unión Europea*. <http://www.anc.org.pe/2017/01/26/proyecto-cso-la2016-377404-union-europea/>. Accessed October 20, 2017.
- ANC. (2018). *Asociación Nacional de Centros*. <https://www.anc.org.pe/>. Accessed November 15, 2018.
- Chianca, T., & Youker, B. W. (2004). Evaluation in Latin America and the Caribbean: An overview of recent developments. *Journal of Multidisciplinary Evaluation*, 1(1), 89–102.
- COEECI. (2018). *COEECI*. <http://www.coeeci.org.pe/quienes-somos/>. Accessed November 15, 2018.
- EvalPerú. (2018a). *Nota sobre la difusión de la conferencia conjunta ReLAC—REDLACME—IDEAS: La evaluación frente a los Objetivos de Desarrollo Sostenible*. <http://evalperu.org/destacados/difusion-de-la-conferencia-con-junta-relac-redlacme-e-ideas-la-evaluacion-frente-los-ods>. Accessed October 20, 2018.
- EvalPerú. (2018b). *Nota sobre el conversatorio sobre estándares de evaluación y políticas públicas*. <http://evalperu.org/destacados/conversatorio-sobre-estandares-de-evaluacion-y-politicas-publicas>. Accessed September 13, 2018.
- EvalPerú. (2018c). *EvalPerú*. <http://www.evalperu.org>. Accessed November 15, 2020.
- ESAN. (2018). *Diploma de Especialización en monitoreo y evaluación de programas y proyectos*. <https://www.esan.edu.pe/diplomas/especializacion-en-monitoreo-y-evaluacion-de-programas-y-proyectos/>. Accessed July 27, 2018.
- La República. (2017). *Transparencia: Coimas de Odebrecht afecta confianza ciudadana en ejercicio de política*. <https://larepublica.pe/politica/847427-transparencia-coimas-de-odebrecht-afecta-confianza-ciudadana-en-ejercicio-de-la-politica>. Accessed December 13, 2018.
- FIDA. (2002). *Gestión orientada al impacto en el desarrollo rural: Guía para el SyE de proyectos*. <https://www.ifad.org/documents/38714182/39726273/intro.pdf/4b3b121d-4e1e-48a9-a124-bc38af04e312>. Accessed July 27, 2018.
- FIDA. (2006). *Conceptos clave de seguimiento y evaluación de programas y proyectos—breve guía*. <http://www.evalperu.org/biblioteca/conceptos-clave-de-seguimiento-y-evaluacion-de-programas-y-proyectos-breve-guia>. Accessed July 27, 2018.
- MCLCP. (2001). *Carta Social: Compromiso por el desarrollo y la superación de la pobreza*. Mesa de Concertación para la Lucha contra la Pobreza.
- MCLCP. (2018). *Mesa de concertación para la lucha contra la pobreza*. <http://www.mesadeconcertacion.org.pe>. Accessed July 27, 2018.

- Ministerio de Desarrollo e Inclusión Social. (2018a). *AYNI Lab Social*. <http://evidencia.midis.gob.pe/innovacion-social/>. Accessed August 15, 2019.
- Ministerio de Desarrollo e Inclusión Social. (2018b). *Informes de evaluación*. <http://evidencia.midis.gob.pe/category/tipo-de-publicacion/informe-de-evaluacion/>. Accessed August 15, 2019.
- Ministerio de Economía y Finanzas. (2015). *Evaluaciones de diseño y ejecución presupuestal (EDEP). Herramienta para el logro de resultados en la gestión pública*. Dirección General de Presupuesto Público.
- Ministerio de Economía y Finanzas. (2017). *Informe de seguimiento del cumplimiento de compromisos de las evaluaciones de diseño y ejecución presupuestal (EDEP) al 31 de diciembre de 2017*. Dirección General de Presupuesto Público.
- Ministerio de Economía y Finanzas. (2018a). *Evaluaciones de diseño y ejecución presupuestal*. <https://www.mef.gob.pe/es/presupuesto-por-resultados/instrumentos/evaluaciones-independientes/655-presupuesto-publico/5356-evaluaciones-de-diseno-y-ejecucion-presupuestal-edep>. Accessed August 15, 2019.
- Ministerio de Economía y Finanzas. (2018b). *Evaluaciones de impacto*. <https://www.mef.gob.pe/es/presupuesto-por-resultados/instrumentos/evaluaciones-independientes/655-presupuesto-publico/5357-evaluaciones-de-impacto>. Accessed August 15, 2019.
- Ministerio de Educación. (2015). Law of organisations and duties. Supreme Decree N° 001-2015-MINEDU. [https://www.peru.gob.pe/docs/PLANES/133/PLAN\\_133\\_2015\\_DS\\_N%C3%82\\_001-2015-MINEDU\(2\).pdf](https://www.peru.gob.pe/docs/PLANES/133/PLAN_133_2015_DS_N%C3%82_001-2015-MINEDU(2).pdf). Accessed August 15, 2019.
- Ministerio de Educación. (2016). *Marco de fundamentación de las pruebas de la Evaluación Censal de Estudiantes*. MINEDU.
- Ministerio de Educación. (2018). *MineduLAB*. <http://www.minedu.gob.pe/minedulab/>. Accessed August 15, 2019.
- On Think Tanks. (2018). *Semana de la Evidencia*. <https://onthinktanks.org/initiatives/semanadelaevidencia/>. Accessed December 20, 2018.
- PUCP. (2017). *Hoja informativa de la diplomatura de monitoreo y evaluación de programas y proyectos*. Not published material.
- PUCP. (2018). *Diplomatura de Monitoreo y evaluación de programas y proyectos sociales*. <http://www.posgradoupch.pe/programa/diplomado-en-monitoreo-y-evaluacion-de-planos-programas-y-proyectos-de-desarrollo/>. Accessed July 27, 2018.
- REDPERUME. (2018). *REDPERUME*. <https://www.mimp.gob.pe/omcp/redperume.php>. Accessed December 13, 2018.
- ReLAC. (2015). *IV conferencia de ReLAC: El futuro de la evaluación en América Latina y El Caribe*. <https://conferenciarelac.wordpress.com/>. Accessed December 13, 2018.

- Shack, N. (2007). *Sistema de monitoreo y evaluación del gasto público a nivel del Gobierno Nacional en Perú*. CLAD (Centro Latinoamericano de Administración para el Desarrollo).
- Shack, N., & Rivera, R. (2017). *Seis años de la gestión para resultados en el Perú (2007–2013)*. Universidad Continental.
- Talledo, M. (2015). Sistema de monitoreo y evaluación en Perú. In G. Pérez & C. Maldonado (Eds.), *Panorama de los sistemas nacionales de monitoreo y evaluación en América Latina* (pp. 311–340). CIDE.
- ULADECH Católica. (2018). *Revistas universitarias peruanas en ScIELO, Scopus, y ESCI/Wos*. <http://reddocente.uladech.edu.pe/profiles/blogs/revistas-universitarias-peruanas-en-scielo-scopus-y-eswi-wos>. Accessed December 13, 2018.
- UPCH. (2018). *Diplomado de monitoreo y evaluación de planes, programas, y proyectos de desarrollo*. <http://www.posgradoupch.pe/programa/diplomado-en-monitoreo-y-evaluacion-de-planes-programas-y-proyectos-de-desarrollo/>. Accessed September 13, 2018.
- USAID. (2017). *Estudio desarrollo de capacidades en proyectos seleccionados de usaid Perú*. Not published material.
- Velásquez, A. (2016). ¿Cómo usar la evidencia para mejorar programas y políticas? In BID, CEPLAN (Ed.). *Los sistemas de monitoreo y evaluación: hacia la mejora continua de la planificación estratégica y la gestión pública* (pp. 98–101).



## Evaluation in the United States of America

*Scott I. Donaldson, Stewart I. Donaldson,  
and Jessica A. Renger*

### INTRODUCTION

Evaluation is now a mainstream activity in American society. The demand for evaluation far exceeds the supply of professional evaluators who are highly trained and hold membership and regularly attend the American Evaluation Association (AEA) and/or other comparable regional or national Voluntary Organisations for Professional Evaluation (VOPEs). The nature of the political environment creates a unique context for evaluation. As a presidential democracy, evaluation activities are often federally mandated and place great emphasis on value constructions and deciding

---

S. I. Donaldson (✉)

Keck School of Medicine, University of Southern California California, Los Angeles, CA, United States

e-mail: [scottdon@usc.edu](mailto:scottdon@usc.edu)

S. I. Donaldson · J. A. Renger

Claremont Graduate University, Claremont, CA, USA

e-mail: [stewart.donaldson@cgu.edu](mailto:stewart.donaldson@cgu.edu)

J. A. Renger

e-mail: [jessica.renger@cgu.edu](mailto:jessica.renger@cgu.edu)



which stakeholders should be involved in discussions about the merit and worth of a programme. While evaluators in the United States (U.S.) do place emphasis on the technical aspects of methodological design and data analysis, there is often a greater focus placed on advocating for social justice and equity than in other countries that operate within a different political context (Greene, 2006).

While evaluation is commonly carried out for projects, programmes, and policies, evaluation activity in the U.S. also includes personnel, product, community, organisational, technology, and systems evaluations among others. These evaluations are often conducted in health, educational, organisational, community, business, and a range of international development contexts. Therefore, it is beyond the scope of any one chapter to adequately describe the wide range of evaluation activity throughout the U.S. The purpose of this chapter is to take on the more modest aim of briefly describing professional evaluation in the U.S.

For the purpose of this chapter, we will define professional evaluation as contracted evaluation services provided by trained evaluators holding memberships in professional evaluation societies and/or their local affiliates like the AEA, which is the largest evaluation professional society in the world with approximately 7,500 members. While evaluations carried out in less formal contexts do constitute a large percentage of evaluation activity in the U.S., capturing this information is beyond the scope of this chapter. It is important to note that some of the latest estimates suggest that there are more than 175 VOPEs and 55,000 members across the world (Donaldson, 2019; Rugh, 2018). The professional evaluators in the AEA belong to 30 regional local affiliates and/or participate in over 60 topical interest groups (TIGs), including government evaluation, advocacy and policy change, business, leadership, and performance, community development, and democracy and governance. Our brief description of professional evaluation in the U.S. will include institutionalising evaluation, current trends in the federal evaluation market, evaluation policy, guidelines for practice, professionalisation, university-based programmes, and evaluation standards.

## INSTITUTIONAL STRUCTURES AND POLITICAL PROCESSES

### *Institutionalising Evaluation in the U.S.*

The U.S. evaluation landscape is in transition from eight years of the Obama administration to a rather different perspective on evaluation and evidence discussed under the Trump administration. There are many perspectives on the advantages and disadvantages of the future of evaluation practice under the new administration. For example, one of the most visible and vocal evaluation theorists and practitioners, Michael Quinn Patton, recently proclaimed that culturally and politically evaluation science is under attack (Patton, 2018). The emergence of antiscience trends in the U.S., such as ‘alternative facts’, ‘fake news’, and a ‘posttruth world’ threaten evidence-based cultures that are sacrosanct to the evaluation community (p. 184). However, now more than ever, evaluation science and evaluation policy decisions can have an impact on the future of domestic policy, national security, and foreign policy outcomes in the U.S.

Notwithstanding Patton’s concern exists a long history of evaluation in the U.S. federal government. Federal agencies such as the Department of Defence, Department of Education, Government Accountability Office, and U.S. Office of Management and Budget have commissioned evaluation services since the 1950s and 1960s with the goal of improving programmatic processes and outcomes. One evaluation policy in particular, the Government Performance and Results Act (GPRA) of 1993 required that all major federal agencies (e.g. The Department of Commerce, The Department of Education, and the Department of Labor) determine their goals, objectives, and a strategy to evaluate results. In 1997, 27 federal agencies submitted reports under the GPRA (Government Accountability Office, 1997). In 2010, the GPRA was modernised to streamline the reporting process, further emphasising federal support for evaluation activities and the continued need for open communication. Of note, although each agency is required to submit an annual report, ensuring high quality data collection and rigorous data analysis is under the jurisdiction of each agency and is not enforced by the federal government. In 2018, the Evidence-Based Policy Making Act was published by congress and required all the U.S. federal agencies to submit an annual systematic plan to address pertinent policymaking questions. As part of this plan, all agencies must develop questions that address key policy issues as well as detail the methods and analytical approaches that will be used

to generate relevant evidence. Additionally, each agency is required to hire “a senior employee as an Evaluation Officer to coordinate evidence-building activities and an official with statistical expertise to advise on statistical policy, techniques, and procedures” (United States Congress, 2018). Thus, while the current U.S. political climate may appear volatile in terms of its attitude towards evaluation, the U.S. federal government continues to commission a high volume of evaluation services and the AEA provides ‘*An Evaluation Roadmap for a More Effective Government*’ (American Evaluation Association, 2016).

The *AEA Evaluation Roadmap* proposed government agencies, policymakers, and programme managers use evaluation to improve programme design, implementation, and effectiveness. The AEA’s vision is to use evaluation as a formative tool to manage government programmes, rather than a reactive strategy to assess programme outcomes. To accomplish this goal, the Executive Branch and Congress need to embrace evaluation as an integral part of good government. However, the evaluation function in government agencies are quite varied in terms of their maturity and breadth. Further, there are varying evaluation needs from federal agencies, such as management, planning, research, policy development, and so forth. As such, the AEA suggested that agencies in the Executive Branch establish evaluation centres and evaluation coordinators to promote evaluation capacity and a system for procuring evaluation services. In tandem, Congress should build connections between evaluations and laws that are passed. For instance, there should be evaluation frameworks within Congress that systematically assess the efficacy of new and continued legislation. Thus, for evaluation to be effective in the U.S. government it is necessary that the Executive Branch and Congress need to take a stance towards using evaluation in authorising legislation. Such a stance will improve existing government programmes and maximise the utility of evaluation services in the U.S. federal government. Although promoted by AEA, many of these changes are yet to be made at the federal level. The Government Accountability Office is one of the few departments that continuously implements evaluations at the request of Congress (American Evaluation Association, 2019). While other agencies such as the Department of Health and the Department of Education have fixed budgets setting aside funds for programme evaluation, widespread and ubiquitous changes to evaluation policy have yet to be realised.

### *Current Trends in the U.S. Federal Evaluation Market*

Lemire et al. (2018a) reviewed recent trends in the U.S. federal evaluation market based on federal evaluation contracts greater than \$3,000, and contracts that related to grants, loans, and other financial assistance. They found that the amount of federal contract funding labelled ‘evaluation’ grew by \$257 million, from \$394 million in 2010, to \$651 million in 2017, from 22 non-defence departments that are subject to the Government Performance and Results Modernisation Act of 2010. On the list of departments examined, the Department of Health and Human Services (DHHS) was responsible for the majority of the contract dollars (\$217 million, or 33.33%), followed by the Department of Veteran Affairs (\$117 million, or 17.97%), and U.S. Agency for International Development (\$74 million, or 11.37%). They also reviewed the procurement and selection processes at the DHHS, which consisted of full and open competition to all prospective clients and a fixed contractual price. Notably, the Government Performance and Results Modernisation Act of 2010 does not differentiate between federal agencies regarding regulations for measures of performance. While the act details key evaluation activities and criteria, it does not specify unique metrics to be used for each federal agency. Thus, prospective clients for the DHHS were responsible for developing appropriate measures of performance that could be feasibly completed given the limitations of a fixed contractual price. Lemire et al.’s analysis revealed there were only eight primary evaluation providers at the DHHS. Lemire et al. (2018b) concluded that while the U.S. Federal Evaluation market has seen an increase in evaluation funding in recent years, there are concerns that the structure of the evaluation market is dominated by a finite number of buyers.

Lemire et al. (2018a) also reviewed the U.S. evaluation market in terms of supply and demand market dynamics that influence evaluation practice. They found that the scope of evaluations conducted ebbs and flows; for instance, there was an increased number of internal evaluations conducted in the 1980s due to increased evaluation capacity and reduced funding from the U.S. federal government. As funding from the federal government increased again in the late 2000s, this trend reversed, and more external evaluations were conducted. Currently the tradeoffs between conducting internal versus external evaluations are being debated among scholars as both are popular approaches in the field. Many scholars cite advantages in organisational learning that are inherent when evaluators are embedded within the organisational structure (Torres & Preskill,

2001), while others advocate for the added value of objectivity external evaluators often possess (Conley-Tyler, 2005). Furthermore, the scope of federal evaluation has broadened to include a focus on both process and outcome-oriented designs. In 2018, the United States Office of Management and Budget outlined key criteria for process and outcome evaluations in the foreign assistance sector, emphasising the necessity of both activities and the importance of matching organisational structure and readiness to the evaluation design (United States Office of Management & Budget, 2018).

In addition, evaluation practice in the U.S. is often conflated with other boundary markets, such as performance management, auditing, and data analytics to name a few. Lemire et al. (2018b) suggested there are multiple evaluation markets and sub-layers within each market. For example, funding for U.S. evaluation projects primarily comes from the public health and education sectors, whereas foundations rarely commission evaluation services. This results in a decentralised evaluation market in the civic sector that is dominated by a few large commissioners. These trends in the U.S. federal evaluation market provide insight into the current landscape of the evaluation market within the federal government, and the importance of evaluation policy in the U.S.

#### *The Importance of Evaluation Policy in the U.S.*

Trochim (2009) outlines several reasons why evaluation policy is important in the U.S. First, it is an important communication mechanism between programme recipients, stakeholders, and organisations. Evaluation policy promotes democratic deliberation in organisations when policies are made explicit and open for dialogue. Preskill (2008) also suggests evaluation policy is a tool for organisational learning because it is open to development and useful for understanding what has worked and not worked in previous circumstances. Finally, evaluation policy has the ability to change evaluation practice. One example is the Programme Assessment system of the U.S. Office of Management and Budget, which required randomised controlled trials as part of the evaluation guidelines. Through numerous discussions on the efficacy of these types of policies, practising evaluators in the U.S., and more specifically, practising evaluators in the AEA have modified past policies and created new evaluation approaches that best serve their clients evaluation needs.

Trochim (2009) developed a taxonomy of evaluation policies that he called '*The Evaluation Policy Wheel*'. It consists of evaluation policy goals,

participation, capacity building, management, roles, process and methods, use, and meta-evaluation. *Goal* policies describe the expected outcomes that are supposed to result from an evaluation. For example, a goal of the evaluation process may be to learn more about the programme or ensure accountability to stakeholders. *Participation* policies refer to the extent to which key stakeholders are involved in the process. *Capacity building, management, and roles* policies involve organisational infrastructure that sponsor and facilitate the utility of evaluation services. *Process* and *methods* policies outline the collection and analysis of data in the organisation (e.g. mixed methods approaches) and ensure security of evaluative data. *Use* policies articulate an action plan for the use of evaluation findings, and recommendations for future evaluations to improve the utility of findings. Lastly, a policy that includes *meta-evaluation* ensures that there is a system to evaluate the evaluations. This helps improve the effectiveness of implementation, quality, and utility of internal evaluations (Trochim, 2009). Although not yet adopted on a federal level, Trochim's Evaluation Policy Wheel has been used as a framework by many evaluators in practice. Most notably, Dillman and Christie (2017) used the policy wheel to analyse the overarching principles guiding evaluations commissioned by the Robert Wood Johnson Foundation. They found that highlighting the importance of evaluation and utilising advisory committees to strengthen evaluation approaches were common policies across all evaluations commissioned. However, the authors also noted many evaluations had policies that changed over time and thus were not always followed consistently or with fidelity.

There are several strengths and limitations to existing evaluation policies in the U.S. For example, in an attempt to make comprehensive evaluation policy, such as described above, there is a threat of over formalisation, which could result in a more bureaucratised environment. There is also the possibility of unintended negative side effects occurring as a result of evaluation policies. Further, Schwandt (2017) calls for a professional ethos of evaluation in the U.S. that is grounded in democratic, co-owned evaluation practice with stakeholders. He argues the technical aspect of evaluation (e.g. competencies, systematic inquiry) has been overemphasised and there is a need for more discussion on moral principles, values, and ways of behaving in evaluation practice that can impact evaluation policy decisions.

### *The AEA Evaluation Policy Task Force*

In 2007, the Board of Directors of the AEA developed an Evaluation Policy Task Force designed to influence evaluation policies that impact evaluation practice. Since then, the primary mission of the task force has been to advise the board of AEA and consult with federal legislators on the state of evaluation policy in the U.S. Although the task force does not have direct decision-making power at a federal level, members are encouraged to advocate for evaluation policy change publicly and consult with staff members in the executive branch (American Evaluation Association, 2009).

While evaluation has the ability to influence policy in all types of political systems (e.g. government, academia, etc.) the purpose of the Evaluation Policy Task Force is to focus on evaluation policy, rather than policy in general. The areas of evaluation policy covered by the task force include: evaluation definition, requirements of evaluation, evaluation methods, human resources regarding evaluation, evaluation budgets, evaluation implementation, and evaluation ethics. For instance, the task force attempts to answer questions such as, how is evaluation designed in each agency, and how does it differentiate itself from related functions such as an audit? The scope of requirements in evaluation may include evaluation procedures and the frequency of commissioned evaluation. Evaluation methods consider which type of methods are required by legislation or regulation, and by which types of initiatives. The task force also considers the training of evaluators and necessary credentials to conduct professional evaluation. They also consider standards for budgeting evaluations, and whether implemented evaluations are guided by specific policies.

### *Guidelines for Evaluation Practice in the U.S.*

The AEA provides the *Guiding Principles for Evaluators* that are intended to facilitate professional ethical conduct of practicing evaluators (American Evaluation Association, 2018). The five principles include systematic inquiry, competence, integrity, respect for people and common good and equity. *Systematic inquiry* ensures that evaluators adhere to the highest technical standards based on evaluator and client time and resources, as well as contextually relevant information to the stakeholders. The *competence* principle ensures that the evaluation team possesses the appropriate knowledge, skills, and abilities necessary to perform professional services

to stakeholders. *Integrity* refers to honesty and transparency between the evaluators and clients relevant to strengths and limitations of the evaluation under consideration. The ability of the evaluators to respect the dignity, well-being, and self-worth of individuals and their cultures is the *respect for people* guiding principle. Finally, *common good and equity* encompasses the purpose of evaluation, which is to contribute a service that benefits society, clients, and recipients of evaluation services. These guiding principles are applied in all sectors that use evaluation in the U.S. for both internal and external evaluations.

Every year, the AEA presents the Alva and Gunnar Myrdal Evaluation Practice Award “to an evaluator who exemplifies outstanding evaluation practice and who has made substantial cumulative contributions to the field of evaluation through the practice of evaluation and whose work is consistent with the ‘AEA Guiding Principles for Evaluators’” (American Evaluation Association, 2018). For example, evaluators who receive this honour have conducted exemplary evaluations, including but not limited to publications, stakeholder feedback, and other forms of evidence that demonstrate the influence of the evaluation work in that particular area. This award provides a template for high-quality, rigorous evaluation practice that serves as a model of evaluation practice in the AEA.

The Centre for Disease Control and Prevention (CDC) also has a practical framework for programme evaluation, which is used as a high-quality template to summarise and organise essential elements of programme evaluation (CDC, 1999). The cycle begins by engaging stakeholders in the evaluation process, including those involved in programme operations (e.g. sponsors, managers, funding officials), those served by the programme (e.g. clients, family members, elected officials), and primary users of the evaluation (e.g. people that will use the evaluation findings). The next involves describing the programme, and the expected outcomes should occur as a result of the programme or intervention. McLaughlin and Jordan (2015) outline the utility of the logic model to help assess programme results and identify ways to improve programme performance by mapping resources, activities, outputs, and outcomes. The third step in the CDC cycle is the focus of the evaluation design that uses time and available resources to suit the information needs of the stakeholders. In this phase it is useful to develop an evaluation strategy and produce useful, feasible, and accurate evaluation findings. It is then the role of the evaluator to gather credible evidence that primary intended users of the programme believe paints an accurate picture of the programme. As a



result of the evidence gathered, the evaluation will justify conclusions on the efficacy of the programme agreed upon by the stakeholders. Finally, it is useful to garner lessons learned from the evaluation process, such as design, preparation, and dissemination of the evaluation findings.

### *Evaluation Use in the U.S.*

The vast number of evaluations conducted in the U.S. and the lack of a universal evaluation quality control and monitoring agency makes it difficult to pinpoint good and bad sectors regarding use of evaluation and evaluation findings. However, Lemire et al. (2018b) suggested public health and education were leading sectors in commissioning evaluation services in the federal government, whereas nonprofit foundations appeared to lag behind. There are several reasons why nonprofits may be more averse to evaluation than federal agencies. Firstly, many nonprofits view evaluation activities as a resource drain on already limited funding and a distraction from their core mission (Carman & Fredricks, 2008). Nonprofits may also fail to see the added value evaluation can provide and view it solely as an external promotional tool instead of a mechanism to improve performance (Carman & Fredricks, 2008). In addition, the funding non-profit organisations receive ranges considerably, bringing with it varying policy implications concerning the use of evaluation. Non-profits that receive federal funding engage in considerably more programme evaluation than nonprofits primarily funded by foundations or local governments (Carman, 2009). The latter group is often not mandated to conduct evaluations and thus more work needs to be done to incentivise non-profits in this sector to incorporate evaluation activities into their regular operations.

Alkin (1985) described that several factors serve as barriers to evaluation use: human factors (e.g. knowledge, skills, and abilities), contextual factors (e.g. political, organisational background), and evaluation factors (e.g. time, budget, ethics). Furthermore, Fleischer and Christie (2009) surveyed 1,140 AEA members on their perceptions, attitudes, and experiences related to evaluation use in theory and practice. They found that AEA members believed stakeholders should be involved in the evaluation process, evaluators should play a number of roles in the evaluation process, and organisational learning is an important outcome of evaluation. Further, the majority of respondents believed a major barrier to use of evaluation was stakeholder rejection of findings based on personal

beliefs rather than data. They also found that the role of internal versus external evaluator influenced the likelihood to plan for evaluation use and organisational outcomes. That is, internal evaluators were more focused on improving the organisational outcomes, whereas external evaluators felt they had to be more objective in their evaluation approach. Many evaluators are starting to take action to promote the use of evaluation in the federal government. In 2017, *EvalAction* an initiative cosponsored by AEA, brought evaluators to Washington D.C. in an attempt to converse with federal policy makers and highlight the importance of using evidence-based decision making in the political process (Washington Evaluators, 2017).

Grob (2018) describes the impact of modern evaluation reports on evaluation practice in the U.S. Whereas, 20 years ago evaluation reports were lengthy technical papers, there has been shift towards electronic reports and slide show (i.e. Powerpoint) presentations. Grob (2018) explained that these slide shows briefs may occur in-person or via teleconferences, include forms of graphic data visualisation, and conclude with discussions with programme stakeholders. He even suggested that high-tech presentations are the de facto of evaluation reporting in the U.S. and are being increasingly disseminated on internet websites.

### *Evaluation in Civil Society*

Evaluators in the U.S. have long recognised the importance of evaluation use not only in the government and for profit sectors, but also in civil society. Civil society has been defined within the evaluation community as the “public that pays attention to social and political issues, the informed public” (Weiss, 1998, p. 28). There has been a call to directly engage those in civil society “to use evaluative information in the program activities in which they are engaged as volunteers, board members, and advisors”, with the realisation that if “informed and active publics know where program shortfalls are and how other programs have overcome them, they may be apt to take a positive stand toward improving the program” (Weiss, 1998, p. 29). Prominent U.S. evaluation theorists have recognised the importance of engaging civil society by advocating for members of the policy shaping community, comprised of “policymakers, civil servants, interested citizens, advocacy groups, and the media, among others” (Greene, 2004, p. 2), to be the target audience for evaluation findings (Cronbach et al., 1980). Henry and Mark (2003) echoed

this sentiment by promoting evaluation be shared broadly within the community and detailed processes for ensuring information is conveyed effectively at the collective level.

Although there has been expressed interest in communicating with civil society through the evaluation dissemination process, there has not been widespread adoption of this idea. One barrier to widespread adoption is that although federal agencies and large foundations dominate the demand side of the U.S. evaluation marketplace, they rarely fund evaluation activities directly. Instead, evaluation services and activities are subcontracted to smaller non-governmental organisations, resulting in “a highly decentralised market for evaluation services in the civic sector” (Lemire et al., 2018b, p. 147). Due to this decentralisation, there is not a singular communication pipeline connecting evaluators working in this sector with interested community members, policy shapers, advocates, and members of the media. Additionally, any evaluation services not directly contracted by the federal government are not required to make data, results, or recommendations accessible to the public. This makes it difficult for those interested to find and use the results of evaluation to make improvements at the local level. Although not federally mandated, large research and evaluation consulting firms often choose to make their findings public. As one of the premier evaluation consulting firms, the RAND organisation regularly publishes the findings and results of their evaluation work to the broader community via their online database (RAND, 2020).

Although there are significant barriers to communication and dissemination within the civil society, efforts have been made by AEA to begin to bridge this gap. AEA has partnered with over 40 non-profit organisations and over 30 local evaluation affiliates with a complimentary mission and a shared set of values to that of the organisation to increase communication between evaluators and organisations across the country (American Evaluation Association, 2020a). Additionally, with the creation of the *Non-profit and Foundation TIG* and the *Use and Influence of Evaluation TIG*, more effort has been extended to find a channel to disseminate evaluation findings to civil society organisations (American Evaluation Association, 2020b).

## PATHWAYS AND STRIDES TOWARDS PROFESSIONAL EVALUATION IN THE U.S.

### *Professionalising a Transdiscipline*

After three decades of vigorous discussions about professionalising evaluation in the U.S. (Altschuld & Engle, 2015), the membership of the AEA voted for the first time to approve a set of general evaluator competencies. This historical decision came about after AEA President Stewart Donaldson established a task force in 2015 to examine whether the AEA membership and U.S. evaluators were ready to agree upon and adopt a set of competencies that evaluation training programme of all types (e.g. degree, certificate, and professional development) could use to train, educate, and develop the next generation of evaluators. This AEA Task Force, spearheaded by Jean King, spent three years engaging evaluators in the U.S. and across the globe to determine the appropriate domains and sub-domains critical for training aspiring evaluators to be competent in their profession. These five domains include:

- **1.0. Professional Domain**—focuses on what makes evaluators distinct as a profession.
- **2.0. Methodology**—focuses on technical aspects on inquiry, such as framing questions, designing studies, sampling, collection, analysing data, interpreting results, and reporting findings.
- **3.0. Context Domain**—focuses on understanding the unique circumstances and settings of evaluations and their users/stakeholders.
- **4.0. Management Domain**—focuses on logistics, such as determining and monitoring work plans, timelines, resources, and other components needed to complete and deliver the study.
- **5.0. Interpersonal Domain**—focuses on human relations and social interactions that ground evaluator effectiveness.

In addition to the new agreed-upon Evaluator Competencies, the AEA provides a set of guiding principles for evaluators (American Evaluation Association, 2018), has provided a statement on cultural competence in evaluation (American Evaluation Association, 2011) and supports the third edition of the ‘*Program Evaluation Standards*’ (Yarbrough et al.,

2011) in an effort to help evaluation training programmes and evaluation practitioners to provide high-quality evaluation services to a wide variety of evaluation clients and consumers. Taken altogether, these efforts towards professionalisation are to further advance practitioner knowledge and skills of how to implement effectively the core knowledge base of the transdiscipline of evaluation.

Evaluation has long been described as a transdiscipline, rather than a typical discipline and profession. For example, Donaldson (2013) and Scriven (1991, 2003) argued that evaluation is one of an elite of group disciplines (e.g. like statistics, design, and logic) that are better described as a transdiscipline. These transdisciplines are notable because they supply tools for other disciplines while retaining an autonomous structure and research effort of their own. For example, there are multiple applications of evaluation practice including the improvement of public health and healthcare, education, psychology and mental health, international development, community development, human resources, organisational development and the like (Donaldson, 2007). Shadish in his 1998 presidential address argues that every transdiscipline and profession, such as evaluation, needs a unique knowledge base. For evaluation, evaluation theory especially supported by research on evaluation provides that unique knowledge base for evaluation professionals today. The challenge that continues to face the U.S. evaluation community is should the attainment of that unique knowledge base by practitioners be formally recognised (e.g. certification, professional designation, accreditation, or licensure) to protect the clients and consumers of evaluation services, as well as ensure evaluation practitioners use the unique knowledge base including the highest ethical standards to guide their practice.

In a recent volume of *'New Directions for Evaluation'* (Altschuld & Engle, 2015), the pathways the U.S. evaluation community might consider following to professionalise the transdiscipline of evaluation are explored in detail. The main pathways that appear on the surface to be feasible include credentialing, certification, and accreditation. Now that the evaluation community has finally reached agreement on what competencies are required to practice evaluation at a high-quality level, it is poised to determine whether these pathways are worth pursuing using these competencies as an agreed-upon foundation. Two other large VOPEs, the Canadian Evaluation Society (CES) and the European Evaluation Society (EES) have already made the decision to provide a professional designation credential for evaluators. While the EES professional

designation programme is relatively new, the CES professional designation programme has been in operation for more than a decade. A recent formative evaluation of the CES professional designation programme provides insights and lessons learned that could be useful for guiding future discussions about professionalisation in the U.S. (Fierro et al., 2016).

### *Recognition of Professional Evaluation in the U.S.*

Recognising evaluation as a profession has been discussed for nearly three decades in the evaluation community. For example, Morell (1990) defined evaluation as a loosely knit group practitioners, while Bickman (1997) in his AEA presidential address asserted evaluation has skills and theories but does not control who enters the profession. He went on to conclude, “we need to move ahead in professionalising evaluation or we will just drift into oblivion” (p. 8). Twenty years have passed since these assertions and the discussion on professionalising evaluation continues to be contentious in the AEA. For instance, in his 2017 presidential address, John Gargani argued that evaluators are like ‘*Ewok*’ forest creatures in the fictional film series Star Wars. Like Ewoks, who are skilled forest survivors, evaluators have a set of skills, knowledge, and abilities that allows them to perform good work. However, both Ewoks and evaluators struggle to shape their core identity in the public sphere, and as a result occupy a niche outside of public awareness.

One potential reason the public fails to recognise evaluation as a profession is the lack of a professional designation. For example, in the U.S., hairstylists, engineers, soccer referees, and all other professions that require training to be a member have more credentials than evaluators. The AEA just recently endorsed an official set of competencies agreed upon by the AEA community. Further, while there is no VOPEs in the AEA like there is in the CES, EES, and United Nations Evaluation Group to name a few, there has been work done to better understand AEA member views on professionalising evaluation. Donaldson (2019) found that AEA member view potential benefits of professionalisation to be improved reputation for the field, trust from consumers of evaluation services (i.e. commissioners, stakeholders, etc.), and an opportunity to assert a clear professional identity. Thus, it is clear AEA members are interested in professionalisation and associated benefits. However, Donaldson

(2019) also found that evaluators are concerned with a potential ‘narrowing effect’, whereby evaluators who have the potential to contribute to the field are alienated based on a set of criteria (e.g. education, experience, competencies). One important finding was that inclusivity and flexibility should be an integral component of a professionalisation process in the U.S. if professionalising evaluation is indeed the path forward. There was also mention of a tiered system (beginner, intermediate, advanced, etc.) in a professionalisation scheme that could organise evaluators based on their skill level. These findings illustrate the importance of embracing diversity (i.e. evaluation approaches, educational background, experience, etc.) when it comes to institutionalising evaluation in the U.S. and including evaluators of all types to enrich the service that evaluation provides.

### *University-Based Programmes (UB) and Professional Development*

Stufflebeam (2001) argues “the evaluation field’s future success is dependent on sound evaluation programme that provide a continuing flow of excellently qualified and motivated evaluators” (p. 445). Lavelle and Donaldson (2010) also share the sentiment that teaching and preparing evaluators is essential for the institutionalisation of evaluation in the U.S. Lavelle and Donaldson (2015) reviewed university-based programmes and professional development opportunities in the U.S. They reported 50 evaluation-specific master’s degrees, 35 certificate programmes, and 40 doctoral programmes focused on preparing future evaluators.

The purpose of university-based formal education in the U.S. is to help participants acquire knowledge, skills, and abilities necessary to be effective practitioners. These programmes were primarily located in departments of education, educational psychology, psychology, and public policy. However, while it is possible practising evaluators in the U.S. have taken university-based training courses, Galport and Azzam (2017) found that less than 10% of the AEA membership receive formal education at a university. Thus, practising evaluators are more likely to receive on-the-job-experience and professional training courses. As training and education are crucial for the establishment of the evaluation profession, Galport and Azzam (2017) conducted a gap-analysis on field-specific competency training-needs, to improve the current training, education, and professional development offered to practising evaluators.

The Claremont Evaluation Centre's (CEC) Workshop Series on Evaluation and Applied Research Methods is the flagship professional development opportunity in the U.S. Workshops can last up to several days and include topics, such as qualitative methods, logic models, mixed methods, data management, et cetera. In addition, the CEC also offers webinars that allow participants to engage in professional development workshops from any location across the world. More recently, the CEC partnered with The Evaluators' Institute (TEI), which is internationally recognised for its high-quality instruction, to deliver a curriculum that emphasises practical knowledge for evaluators (Claremont Evaluation Centre 2018). TEI offers professional training courses, evaluation certificates, and corporate education in various locations across the country (e.g. DC Metro Area, Claremont and Chicago). The TEI faculty represent some of the most well-renowned evaluation theorists and practitioners in the world, including Michael Quinn Patton, Mark Lipsey, and Stewart Donaldson, among many more.

### *Profession/Discipline*

Several universities professorships and professional journals exist in the U.S. For example, in the psychology department at Claremont Graduate University, there are Full Professors of Evaluation and Applied Research Methods, an Associate Professor of Evaluation and Applied Research Methods, an Assistant Clinical Professor of Evaluation, and a Research Associate Professor of Evaluation. The University of California, Los Angeles houses evaluation professors in the department of education. The department of Organisational Leadership, Policy, and Development at the University of Minnesota has evaluation professors. Finally, Western Michigan University and University of Connecticut have Professors in Evaluation, Measurement, and Research. Major journals in the field include, *The American Journal of Evaluation*, *Evaluation*, and *Evaluation and Program Planning*. *The American Journal of Evaluation*, *Evaluation* and *Evaluation and Program Planning* have been publishing since 1981, 1995, and 2001 respectively, all producing four issues annually, each containing approximately ten articles. It is important to note that due to the transdisciplinary nature of the evaluation field, many practice-based articles are published in substantive areas (i.e. health and education journals) rather than in journals specific to the field of evaluation.



### *Compliance to Standards and Quality Obligations*

As aforementioned, there is no professionalisation system (e.g. credentialing, certification, accreditations) that currently exists in the AEA. However, McDavid and Huse (2015) outlined key terms and concepts to professionalise individuals and institutions that are possible avenues the AEA could consider. There are *Ethical Guiding Principles* that reflect the core values of the AEA and guide professional ethical conduct of evaluation (American Evaluation Association, 2018). The five guiding principles include systematic inquiry, competence, integrity, respect for people, and common good and equity (See AEA Guiding Principles for definitions). Other efforts towards professionalisation include Montrosse-Moorhead and Griffith's (2017) Checklist for Evaluation-Specific Standards, which operate as standards framework that can be reported in evaluations (e.g. contextual information, systematic inquiry activities, and stakeholder involvement). To our knowledge, there are few attempts by clients of evaluation services to demand credentials for evaluation services, nor do evaluators themselves utilise the AEA Guiding Principles. In this regard, there is much progress that needs to be made towards institutionalising standard and quality controls.

## CONCLUSION

Evaluation practice in the U.S. is vast and robust. Evaluation is going on in just about every sector of our society. However, we believe much of this evaluation practice is not informed by the knowledge base of professional evaluation. It is difficult to accurately describe evaluation being conducted by researchers and 'evaluators' not trained in evaluation theory and practice, or part of the professional evaluation community. The only way we felt we could provide an accurate picture of the U.S., perhaps the largest or one of the largest evaluation communities across the globe, was to focus where we have concrete data that accurately describe evaluation practice—the professional evaluation community.

The purpose of this chapter was to briefly describe the range of activities professional evaluators working in the U.S. often engage in. It is important to point out that this is a rather small slice of all the evaluation activity and services being provided to government agencies, educational

institutions, health systems and organisations, philanthropic communities, non-profit and for-profit businesses, and a vast array of organisations supporting international development initiatives. The breadth and complexity of the U.S. federal government and civil society make synthesising the wide range of evaluation efforts taking place difficult. Although there is a vast room for improvement regarding the professionalisation of evaluation in the U.S., AEA's adoption of evaluator core competencies and guiding principles as well as the increasing number of university-based training programmes represent promising starts towards formalising the discipline and profession. We hope the sample of topics and activities described in this chapter provides a useful snapshot of professional evaluation in the U.S. today, as well as highlights that evaluation activity more generally is much more prevalent and robust than in most other regions of the world.

### LIST OF ABBREVIATIONS

AEA	American Evaluation Association
CDC	Centers for Disease Control
CEC	Claremont Evaluation Center
CES	Canadian Evaluation Society
DHHS	Department of Health and Human Services
GPRA	Government Performance and Results Act
EES	European Evaluation Society
TEI	The Evaluators Institute
TIGs	Topical Interest Groups
VOPEs	Voluntary Organisations for Professional Evaluation
U.S.	United States

### REFERENCES

- Alkin, M. C. (1985). *A guide for evaluation decision makers*. Sage.
- Altschuld, J. W., & Engle, M. (2015). The inexorable historical press of the developing evaluation profession. In J. W. Altschuld & M. Engle (Eds.), *Accreditation, certification, and credentialing: Relevant concerns for U.S. evaluators: New directions for evaluation* (pp. 5–19). Jossey-Bass.

- American Evaluation Association. (2009, July). *AEA evaluation policy task force charge*. <https://www.eval.org/p/cm/ld/fid=151>. Accessed on 7 March 2019.
- American Evaluation Association. (2011). *American evaluation association public statement on cultural competence in evaluation*. Retrieved from [www.eval.org](http://www.eval.org). Accessed on 11 November 2020.
- American Evaluation Association. (2016, September). *An evaluation roadmap for a more effective government*. <http://www.eval.org/d/do/4008>. Accessed on 9 March 2020.
- American Evaluation Association (2018, July). *Guiding principles for evaluators*. <https://www.eval.org/p/cm/ld/fid=51>. Accessed on 7 March 2020.
- American Evaluation Association. (2019). *An evaluation roadmap for a more effective government*. <https://www.eval.org/evaluationroadmap>. Accessed on 9 March 2020.
- American Evaluation Association. (2020a). *Professional groups*. <https://www.eval.org/p/cm/ld/fid=11>. Accessed on 7 March 2020.
- American Evaluation Association. (2020b). *Topical interest groups*. <https://www.eval.org/p/cm/ld/fid=11>. Accessed on 9 March 2020.
- Bickman, L. (1997). Evaluating evaluation: Where do we go from here? *Evaluation Practice*, 18, 1–16.
- Carman, J. G. (2009). Nonprofits, funders, and evaluation: Accountability in action. *The American Review of Public Administration*, 39(4), 374–390.
- Carman, J. G., & Fredericks, K. A. (2008). Nonprofits and evaluation: Empirical evidence from the field. *New Directions for Evaluation*, 2008(119), 51–71.
- Centers for Disease Control and Prevention. (1999). Framework for program evaluation in public health. *Morbidity and Mortality Weekly Report*, 48(11), 1–42.
- Claremont Evaluation Center. (2018). *The evaluators' institute*. <https://teicgu.edu/>. Accessed on 3 March 2020.
- Conley-Tyler, M. (2005). A fundamental choice: Internal or external evaluation? *Evaluation Journal of Australasia*, 4(1), 3–11.
- Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., Walker, D. F., & Weiner, S. S. (1980). *Toward reform of program evaluation* (p. 3). Jossey-Bass.
- Dillman, L. M., & Christie, C. A. (2017). Evaluation policy in a nonprofit foundation: A case study exploration of the Robert Wood Johnson Foundation. *American Journal of Evaluation*, 38(1), 60–79.
- Donaldson, S. I. (2007). *Program theory-driven evaluation science: Strategies and applications*. Psychology Press.
- Donaldson, S. I. (Ed.). (2013). *The future of evaluation in society: A tribute to Michael Scriven*. Information Age Publishing, INC.

- Donaldson, S. I. (2019). Where do we stand? Recent AEA member views on professionalization. *Evaluation and Program Planning*, 72, 152–161.
- Donaldson, S. I., & Donaldson, S. I. (2016). Visions for using evaluation to develop more equitable societies. In S. I. Donaldson, & R. Picciotto (Eds.), *Evaluation for an equitable society*. Information Age Publishing, Inc.
- Fierro, L. A., Galport, N., Hunt, A., Codd, H., & Donaldson, S. I. (2016). *Canadian evaluation society credentialed evaluator designation program*. Claremont Evaluation center technical report.
- Fleischer, D. N., & Christie, C. A. (2009). Evaluation use: results from a survey of U.S. american evaluation association members. *American Journal of Evaluation*, 30(2), 158–175.
- Galport, N., & Azzam, T. (2017). Evaluator training needs and competencies: A gap analysis. *American Journal of Evaluation*, 38(1), 80–100.
- Government Accountability Office. (1997). *Reports on the government performance and results act*. <https://www.gao.gov/new.items/gpra/gpra.html>. Accessed on 4 March 2020.
- Greene, J. C. (2004). The educative evaluator: An interpretation of Lee J. Cronbach's vision of evaluation. *Evaluation roots: Tracing theorists' views and influences* (pp. 169–180). Sage.
- Greene, J. C. (2006). Evaluation, democracy, and social change. *The Sage handbook of evaluation* (pp. 118–140). Sage.
- Grob, G. (2018). Evaluation practice: Proof, truth, client relationships, and professional growth. *American Journal of Evaluation*, 39(1), 123–132.
- Henry, G. T., & Mark, M. M. (2003). Beyond use: Understanding evaluation's influence on attitudes and actions. *American Journal of Evaluation*, 24(3), 293–314.
- LaVelle, J., & Donaldson, S. (2010). University-based evaluation training programs in the United States 1980–2008: An empirical examination. *American Journal of Evaluation*, 31(1), 9–23.
- LaVelle, J., & Donaldson, S. (2015). The state of preparing evaluators. *New Directions for Evaluation*, (145), 39–52.
- Lemire, S., Fierro, L. A., Kinarsky, A. R., Fujita-Conrads, E., & Christie, C. A. (2018a). The U.S. federal evaluation market. In S. B. Nielsen, S. Lemire, & C. A. Christie (Eds.), *The evaluation marketplace: Exploring the evaluation industry*. *New Directions for Evaluation*, 160, 63–80.
- Lemire, S., Nielsen, S. B., & Christie, C. A. (2018b). Toward understanding the evaluation market and its industry—Advancing a research agenda. In S. B. Nielsen, S. Lemire, & C. A. Christie (Eds.), *The evaluation marketplace: Exploring the evaluation industry*. *New Directions for Evaluation*, 160, 145–163.
- McDavid, J. C., & Huse, I. (2015). How does accreditation fit into the picture? *New Directions for Evaluation*, 2015(145), 53–69.

- McLaughlin, J., & Jordan, G. (2015). Handbook of practical program evaluation. In J. S. Wholey, P. H. Harry, & K. E. Newcomer (Eds.), *Using logic models* (pp. 62–87). Wiley.
- Montrosse-Moorhead, B., & Griffith, J. (2017). Toward the development of reporting standards for evaluations. *American Journal of Evaluation*, 38(4), 577–602.
- Morell, J. (1990). Evaluation: Status of a loose coalition. *Evaluation Practice*, 11(3), 213–219.
- Patton, M. Q. (2018). Evaluation science. *American Journal of Evaluation*, 39(2), 183–200.
- Preskill, H. (2008). Evaluation's second act: A spotlight on learning. *American Journal of Evaluation*, 29, 127–138.
- RAND. (2020). *Research: Informing the current public policy debate and topics on the public agenda*. <https://www.rand.org/research.html>. Accessed on 5 March 2020.
- Rugh, J. (2018, September 23). What Are VOPEs? And How Do They Relate To Us As Evaluators? <https://aca365.org/blog/what-are-vopes-and-how-do-they-relate-to-us-as-evaluators/>
- Schwandt, T. (2017). Professionalization, ethics, and fidelity to an evaluation ethos. *American Journal of Evaluation*, 38(4), 1–8.
- Scriven, M. (1991). *Evaluation thesaurus*. Sage.
- Scriven, M. (2003). Evaluation theory and metatheory. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 15–30). Springer.
- Stufflebeam, D. (2001). Evaluation models. *New Directions for Evaluation*, 89, 7–98.
- Torres, R. T., & Preskill, H. (2001). Evaluation and organizational learning: Past, present, and future. *American Journal of Evaluation*, 22(3), 387–395.
- Trochim, W. (2009). Evaluation policy and evaluation practice. *New Directions for Evaluation*, 123, 13–32.
- United States Congress. (2018). *H.R.4174—Foundations for evidence-based policymaking act of 2018*. <https://www.congress.gov/bill/115th-congress/house-bill/4174>. Accessed on 9 March 2020.
- United States Office of Management and Budget. (2018). *Monitoring and evaluation guidelines for federal departments and agencies that administer United States foreign assistance*. <https://www.whitehouse.gov/wp-content/uploads/2017/11/M-18-04-Final.pdf>. Accessed on 8 March 2020.
- Washington Evaluators. (2017). *Eval action: Evaluators visiting with policymakers about the importance of evaluation*. <https://washingtonevaluators.org/event-2456751>. Accessed on 9 March 2020.
- Weiss, C. H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19(1), 21–33.

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Sage.

PART III

---

# Transnational Organisations and Networks



## CLEAR LAC—Centers for Learning on Evaluation and Results—Latin America and the Caribbean

*Claudia Maldonado Trujillo*

“CLEAR is a global network dedicated to improving policy, planning, and implementation through strengthening monitoring and evaluation systems and capacities. We innovate and learn locally and regionally, and share and inspire globally”.<sup>1</sup> This chapter analyses the case of the Regional Center for Learning on Evaluation and Results for Latin America and the Caribbean (CLEAR LAC, 2012–2019) as an example of a new generation of capacity-building efforts in the Global South and its linkages with the institutionalisation of evaluation. I argue that the transformative vision of the CLEAR initiative has been particularly suitable as a capacity-building strategy in the Latin American context. In this comparatively favourable environment, the basic tenets of the vision behind the CLEAR initiative

---

<sup>1</sup> Mission statement of the CLEAR initiative (<http://www.clearinitiative.org>).

---

C. Maldonado Trujillo (✉)  
Autonomous Metropolitan University (UAM), National Council for the  
Evaluation of Social Development Policy (CONEVAL), Mexico city, Mexico  
e-mail: [cmaldonado@correo.xoc.uam.mx](mailto:cmaldonado@correo.xoc.uam.mx)



seem to have held. At the same time, however, the global component of the initiative has developed more unevenly and may take more time and additional financial and human resources to achieve its full potential. Finally, I present an outlook of the main strengths and limitations of this type of capacity-building strategy, with a special focus on recent political trends in Latin America.

In the first section, I lay out the origins of CLEAR as a capacity-building strategy in the context of recent changes in global discourse on development and development effectiveness. I argue that this initiative represents a new vision for evaluation capacity building (ECB) for its emphasis on ownership and South-South learning as a mechanism to promote good government through the institutionalisation of monitoring, evaluation and performance-based management. In other words, as an anchor of results-based mutual accountability. In the second section I describe the evolution of the organisational structure and transformative vision of CLEAR as a global initiative. The third section I review the trajectory of CLEAR LAC as a member of the global initiative, and show how the Centre's configuration and strategy can reasonably claim contributions to institutionalisation and capacity-building in various dimensions.<sup>2</sup> Finally, the fourth section reflects on the strengths and weaknesses of this strategy as well as some implications of the Latin American political climate with respect to prospects for continued strengthening and institutionalisation of evaluation.

## THE CLEAR INITIATIVE: RETHINKING CAPACITY-BUILDING FOR DEVELOPMENT

The CLEAR initiative was launched in 2010 as a multi-donor initiative originally designed to promote capacity-building in monitoring, evaluation and performance-based management in order to improve decision-making in the developing countries. The emergence of this initiative is best understood in the context of a new narrative on international cooperation and development effectiveness, and the emergence

<sup>2</sup> Due to space limitations, I do not analyse the case of CLEAR Brazil because it is the Centre that was established at last and it has a slightly different definition of regional scope, including work in Lusophone Africa. The host institution and the enabling environment in Brazil are comparable to CLEAR LACs, but I am not in position to assess what the African context entails for this Centre.

of a more bottom-up, horizontal notion of capacity-building in response to widespread discredit of donor-driven development cooperation and its detrimental effects on development progress.

The public recognition of the need to rethink development effectiveness and the governance structure required to ensure transparency and accountability was clearly expressed in the Paris Declaration (2005). In contrast with previous international statements on cooperation, this was the first shared commitment by donor and aid recipients (partners) to adopting specific measures to improve development effectiveness with the adoption of the following principles: ownership, alignment, harmonisation, results and mutual accountability. These principles entailed the adoption of a results-based framework for development, the recognition that developing countries could establish their own strategies for poverty reduction, institutional development and fighting corruption, that donors would follow these strategies, operate through local systems rather than build externally driven parallel development interventions and prevent duplication and fragmentation by sharing information and enhancing coordination and framing international cooperation as a mutually accountable system between donors and partners. The Paris Declaration was followed by the Accra Agenda for Action in 2009, which further expanded the notion of ownership and inclusiveness by emphasising coordination, leadership and participation of developing countries, as well as a focus on ‘real and measurable impact’ and capacity development, broadly defined as the ability of countries to manage their own future. In 2011, the Busan Alliance for Development Cooperation Effectiveness further emphasised the need to balance shared responsibility with technical assistance for national capacity development which could then enable better implementation (through monitoring and evaluation). Furthermore, the changing ecosystem of international cooperation was explicitly recognised: multiple actors (national, subnational, non-governmental), multiple channels for cooperation (South-South, notably) and within and across sector heterogeneity.

These new pillars of international development policy discourse were perceived as necessary catalysts of development cooperation in order to meet the 2015 target set by the United Nations for the Millennium Development Goals. They were soon incorporated into the institutional language of multilateral development banks. The logic is rather straightforward: if developing countries are expected to fulfil an equal partner role in a mutual accountability framework, these countries are expected to

develop their own capacities for results-oriented implementation in order to effectively channel aid through their local systems and exercise real ownership as a mean to sustainability. With these considerations in mind, it is evident that the CLEAR initiative's original formulation is a direct derivation of this new development thinking: (1) it is primarily focused on system capacity development in M&E; (2) it is focused on the Global South to promote ownership and (3) it is conceived as a learning platform in the shape of a global network of regionally based Centres.<sup>3</sup>

The starting premise of CLEAR is that capacity cannot be just transferred as financial resources, but that it needs to be context-specific to be organically developed and viable in the long term (persist after the technical assistance support is gone). Furthermore, development institutions started to envision new models for capacity-building that entailed multi-stakeholder and multidimensional interventions rather than *just training* efforts:

Efforts to improve support for capacity development have been mixed. While donors met the target on co-ordinated technical co-operation, support for capacity development often remains supply-driven, rather than responding to developing countries' needs. (OECD, 2012, p. 18)

Capacity development had evolved from the emphasis on training individuals to the more nuanced notion of capacity development that acknowledges that "organisations work within an enabling environment that consists of policies, networks and an attitude to engagement" (OECD, 2012, p. 3). This conceptual and strategic stance clearly reflects recent thinking on institutional design and development in the social sciences, as well as the growing emphasis on the role of the enabling environment and meta-institutional background conditions for formal norms, practices and routines to be effectively internalised and reproduced. This, in a nutshell, reflects many of the core features behind CLEAR.

<sup>3</sup> According to programme documents, "CLEAR is a global M&E capacity development programme that brings together academic institutions and donor partners to foster the collection, measurement, analysis, and subsequent use of robust evidence in developing countries' policy and programmatic decision making" (<http://www.theclearinitiative.org>).

## THE CLEAR INITIATIVE: STRUCTURE, GOVERNANCE AND VISION

The CLEAR initiative is an ambitious capacity-building effort anchored in the global south that seeks to create a multi-level M&E capacity-building network by directly incubating regional nodes in well-established institutions with regional relevance and providing for knowledge generation and lesson-exchange opportunities at the global level. The CLEAR initiative is a global network with a Global Hub/Secretariat (hosted at the Independent Evaluation Group (IEG) of the World Bank) that sponsors regional Centres for Learning on Evaluation and Results that are nested in prestigious academic institutions and/or academic consortia with regional presence and scope. Centres' host institutions were selected after a carefully designed competitive process that included site visits, conversations with stakeholders and a thorough review of the institutions' vision for CLEAR (a regional demand-assessment and a regional strategy) as well as proven institutional strengths and elective affinity with CLEAR's core objectives.

As a global network, the main expectation of CLEAR is that, in contrast with previous capacity-building strategies, CLEAR Centres would enable ownership, sensitivity to context and cultural pertinence of capacity-building efforts due to the fact that they would be embedded in their local and regional policy dialogues, and would thus have access to better information on the demand structure and the changing environment that shapes the policy space for results-oriented tools and the promotion of an evaluation culture. The need for embeddedness of these efforts greatly influenced the initiative's design. CLEAR Centres are expected to work, learn and exert influence on M&E and PBM systems from the region and for the region. Centres are also expected to showcase and package their work and experience as learning and knowledge products that could in turn fuel inter-regional cross-fertilisation and global learning by extension. Complementarities and synergies among centres were also expected to develop over time, once the region-specific identity and strengths were jointly developed with stakeholders and shared between centres directly and through the Global Hub.

Host institutions were academic institutions with proven expertise in M&E, policy research and, more generally, the promotion of evidence-informed policy-making. This would secure access to a reliable stock of human capital to provide methodologically sound assistance, on the one hand, and the political and technical credibility and convening

power on the other. In principle, strong and independent academic institutions can exert some influence (soft power) on governmental and non-governmental stakeholders through training, research publications and professional networks in a wide range of institutional arenas (multilateral organisations, governments, regional networks, etc.). This design principle, namely to nest centres in academic institutions, also came with the expectation that, these centres would have methodological and political credibility but also the incentives and resources to actively engage with government, civil society and attentive publics in a more proactive, timely and service-oriented manner.<sup>4</sup>

To date, there are six regional CLEAR Centres focused on Anglophone Africa (University of Witswatersrand's in Johannesburg, South Africa), Brazil and Lusophone Africa (Fundação Getulio Vargas in São Paulo, Brazil), Francophone Africa (Centre Africain d'Etudes Supérieures en Gestion in Dakar, Senegal), Latin America (Center for Research and Teaching in Economics in Mexico-city, Mexico), East Asia (Shanghai National Accounting Institute, in China) and South Asia (Institute for Financial Management and Research in India, Fig. 13.1).



**Fig. 13.1** Global scope of the CLEAR regional centres and global hub (<http://www.theclearinitiative.org>)

<sup>4</sup> Host institutions raised to the challenge, but expectedly faced some tensions between meeting traditional academic standards (double-blind peer review publications) and the advocacy and knowledge dissemination dimensions of CLEAR.

The original group of donors (African Development Bank, Asian Development Bank, DFID, SIDA, Rockefeller Foundation, Inter-American Development Bank and the World Bank) were represented in the Board through their respective heads of evaluation. This is a distinctive feature of CLEAR, that it is championed by those directly involved in the evaluation function and capacity development within donor agencies. In my view, this is a unique feature of CLEAR that potentially provided Centres with privileged access to knowledge, professional networks and a shared ethos regarding the relevance of the evaluation function and the challenges of the field, in the field. Furthermore, this made the CLEAR initiative more responsive to windows of opportunity, country and regional level technical assistance needs and indirectly to internal debates on evaluation capacity development and the changing perspectives among donor agencies on the need to improve, for instance, knowledge management and usability over standard independence and credibility criteria in evaluation.

The governance structure of CLEAR evolved with it. Over time, a very general South-South, (yet still donor-driven) vision for CLEAR became a more nuanced, diverse and complex theory of change that was co-designed and conceptualised by the Centres. Centres had themselves learned and evolved due to their interaction with regional partners and stakeholders. The IEG at the World Bank served as Secretariat, donor partners as Board members and the Centres adopted the internal governance structure compatible with their host institution, provided that accountability mechanisms such as regional advisory boards were in place. Some Centres appointed academics and professors as heads of the Centres while others appointed programme officers with managerial experience on M&E to lead these efforts. Likewise, the internal structure and operational configuration of Centres were expected to capture region-specific needs and service provision comparative advantages.

Overall, CLEAR has gone through two phases of funding and strategic reformulation and will enter a third phase in 2020. Individual Centres are, however, at a different time-horizon and maturation stages and their respective grant designs vary accordingly.<sup>5</sup> Another key expectation from Centres is that they would become self-sustainable in the medium term

<sup>5</sup> CLEAR Centres have access to different amount of funding from different donors. For this reason, Centres grants, target and strategies vary by region and source of funding. As an example, most of the international funding for the CLEAR LAC was provided by

through the development of successful business models that secured a steady source of revenues from user fees, grants and/or non-profit funding.<sup>6</sup>

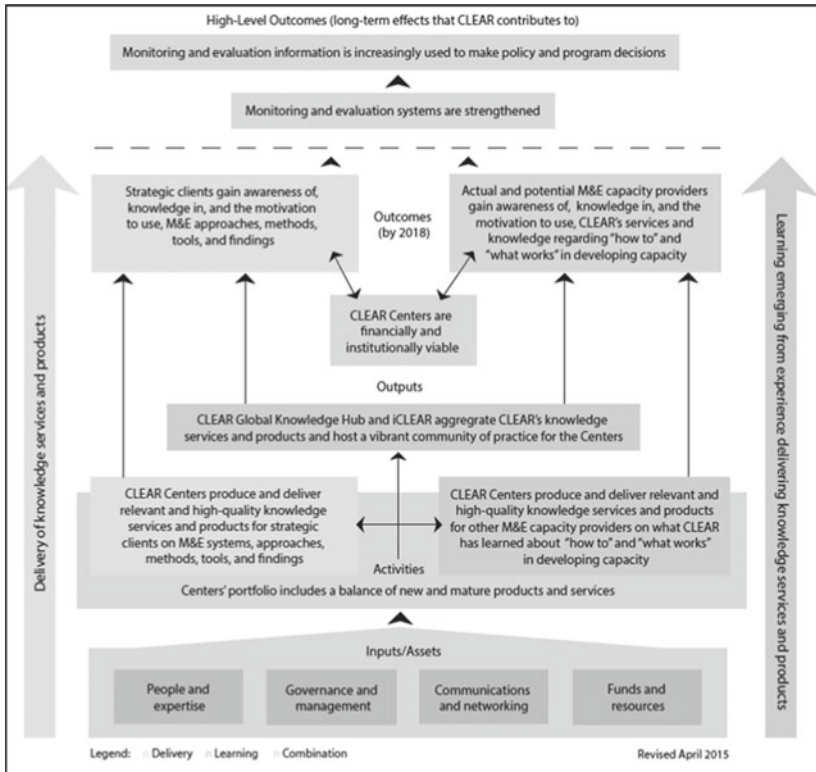
At the beginning of phase II, the initiative underwent some important adjustments. First of all, in terms of global governance, the CLEAR initiative started to devolve power and voice to the Centres in global governing and decision-making bodies, as well as strengthen the responsibility of the global hub as a network enabler. The current governance structure includes a Council (comprised by Centre Directors, Funding Committee Members and the CLEAR CEO). The Funding Committee represent the donors, including a senior manager from IEG at the World Bank and guides funding strategy and accountability. The Global Hub is still located at IEG. In addition to governance, CLEAR revised its strategic outlook. Through collaborative work with members and stakeholders, CLEAR Centres explicitly articulated a revised Theory of Change for the initiative that included a new emphasis on knowledge products and innovation. Among other factors, the revised Theory of Change responded to an external evaluation finding that suggested that despite having been designed as a pilot experimenting with a new outlook on capacity-building, CLEAR had not been designed to be able to learn from implementation and monitor progress and eventual impacts. In addition to this, internal discussions and external pressures to further articulate the distinctiveness of CLEAR as a model for learning and innovation, provided the basis for a revised Theory of Change and the adoption of Key Performance Indicators in 2015 (Fig. 13.2).

CLEAR Centres generally provide training services, technical assistance and knowledge generation and dissemination services with a region-specific mix and locally differentiated sectoral and thematic foci. They typically engage with national and subnational government agencies and legislative bodies by providing advice, training services and knowledge exchange,<sup>7</sup> as well as with civil society and professional networks in order to further advance the evaluation culture, mobilise evaluation

the Inter-American Development Bank. During phase I, the Mexican government matched this funding to support CLEAR.

<sup>6</sup> According to the Mid Term Evaluation, the Centre for Latin America was the only one with real probabilities of sustainability after phase II.

<sup>7</sup> In 2018, CLEAR delivered capacity-building training to almost 30,000 individuals from 65 countries.



**Fig. 13.2** CLEAR theory of change phase II (<http://www.theclearinitiative.org>)

results for public debate and generally socialise the quest for better use of evidence for improved decision-making.<sup>8</sup> These trainings were often accompanied with technical assistance services to improve M&E systems and promote evidence-informed decision-making. As a global network, CLEAR launched PRIME (Programme in Rural M&E) in 2017 as a large-scale capacity-building and certification framework programme that will professionalise M&E in the rural development sector with basic to

<sup>8</sup> More detailed information on the centers' activities and the Global Hub at <http://theclearinitiative.org>.



advanced courses delivered in Spanish, English, French—and eventually Portuguese.<sup>9</sup> This capacity-building service is the first joint programme that involves the participation of all the centres and is a promising consolidation project for CLEAR as a Global Network.

Today, as partners and stakeholders reflect on what a phase III for CLEAR would entail and how Centres should collaborate with similarly oriented capacity-building efforts in their respective regions and on a global scale, the long-term sustainability of these efforts is still an open question. As I will argue in the case of Latin America, the viability and continuity of these efforts are closely dependent on national and regional demand-side factors that are exogenous to CLEAR for the most part, yet significantly affect its prospects to contribute to the institutionalisation of monitoring, evaluation and performance-based management. So far, the Centres have made remarkable progress in their consolidation and have been clearly able to influence institutionalisation processes, strengthen regional linkages, formal and informal networks and promote an evaluation culture.

### THE CLEAR CENTER FOR LATIN AMERICA AND THE CARIBBEAN

We strengthen M&E capacity across Latin America and the Caribbean through training courses and workshops for international organisations, government, civil society, academia, and the private sector. We also share knowledge to promote evidence-based decision making and improve policy and programmes in the region.<sup>10</sup>

In 2011, the Centre for Research and Teaching in Economics (CIDE)<sup>11</sup> was competitively selected as the host for the CLEAR Center for Latin America and the Caribbean from a pool of 22 applications from seven Latin American Countries, including the regional leaders on M&E. At

<sup>9</sup> Through PRIME, CLEAR has trained 200 people working in the rural development field from 75 countries.

<sup>10</sup> Source: <http://theclearinitiative.org>.

<sup>11</sup> CIDE is a leading public research institution specialised in the social sciences founded in 1974.

that time, the Mexican Government was actively promoting performance-based management, monitoring and evaluation at the federal and subnational levels due to a series of legislative mandates enacted in the preceding years that sought to institutionalise results-orientation in public administration.<sup>12</sup> For this reason, the Mexican government offered to match CLEAR's funding during the first four years of implementation (2012–2016).<sup>13</sup> The Centres' organisation was structured around three types of services: training, technical assistance and knowledge products with the overall strategy to create linkages and outreach at a regional scale. As a launching strategy, the Centre focused on actively engaging with institutions and networks that were already regional in scope and vocation,<sup>14</sup> advancing collaboration and promoting dialogue with other academic institutions promoting evidence-based policy<sup>15</sup> and creating linkages with civil society actors (such as CIPPEC in Argentina, GRADE in Perú and México Evalúa in Mexico), as well as partner initiatives, like FOCEVAL in Costa Rica and UNDP's National Evaluation Capacity in the global South (NEC). This web of collaboration facilitated direct contact with governments and high-level public officials in countries that had showed interest in advancing their M&E systems and building human capital for them (e.g. Argentina, Peru, El Salvador, Costa Rica, Uruguay and Paraguay), as well as countries that were already championing this agenda, like Chile and Colombia. The Centre aspired to become a full-service resource centre for different degrees and levels of training,

<sup>12</sup> For a detailed account of this process, see Pérez, Y., Maldonado, C., & Faustino, D. G. (2015). El Sistema de seguimiento y evaluación de programas federales en México: Retos para su consolidación. In G. Pérez, & C. Maldonado (Eds.), *Panorama de los sistemas nacionales de monitoreo y evaluación en América Latina*, (pp. 273–311). México: CIDE.

<sup>13</sup> The Centre was formally inaugurated by the President in the Presidential Palace (Los Pinos), with a high-level delegation of government officials, Latin American ambassadors in Mexico, civil society, academia, the private sector and the head of evaluation of IDB representing the donors.

<sup>14</sup> Such as the Latin American Centre of Administration for Development (Centro Latinoamericano de Administración para el Desarrollo, CLAD), the Organisation of American States (OEA) and M&E regional networks such as REDLACME and RELAC.

<sup>15</sup> Such as the Universidad Católica de Chile and Universidad Nacional del Lanús in Argentina, Pontificia Universidad Javeriana in Colombia and national institutes for public administration like ESAP in Colombia and INAP in Brazil.

technical assistance and knowledge product generation by cultivating an emerging reputation of active partnership and high-quality services:

“Delivering custom training courses and workshops to help leaders and practitioners working in government, international organisations, civil society, and the private sector strengthen their M&E knowledge and skills. Providing rigorous, practical evaluations and technical assistance, and conducting applied research for public, private, and non-profit organisations. Producing and sharing evidence-based knowledge products to strengthen M&E in Latin America and the Caribbean. Collaborating with regional M&E partners to organize Evaluation Week and other knowledge-sharing events.”<sup>16</sup>

By 2019, the Centre had a team of 20 full-time staff,<sup>17</sup> backed by a network of over 50 M&E experts from Latin America and abroad (academics and consultants). It had successfully delivered 60 courses in 14 Latin American countries, rated as high-quality courses by participants and directly trained over 2.000 people from over 20 countries in the region, performed 20 evaluations/technical assistance projects and published over a dozen books and knowledge products in Spanish documenting the Latin American experience on M&E as well as providing learning tools, concepts and methods for evaluation in the local language.

The CLEAR centre was directly involved in the discussion on national evaluation policies and strategies in Mexico, Argentina, Costa Rica, Perú, El Salvador and St Lucia, and facilitated dialogue between high-level political stakeholders from all the Central American Countries (Roundtable on M&E in Guatemala, 2015). Furthermore, it actively engaged with civil society efforts at mobilising evidence-based decision-making and results-oriented accountability and trained a significant number of public servants directly involved in evaluation practice in their respective countries. The Centre also provided technical assistance on measuring impact and the adoption of results-frameworks and indicators for entrepreneurs and emerging evaluators, and promoted an innovative internship and graduate research award to promote emerging M&E specialists from the region. It sponsored pilots for projects for gender-sensitive training projects, as well as facilitated the organisation and

<sup>16</sup> Source: CLEAR LAC at <http://theclearinitiative.org>.

<sup>17</sup> The Centre was launched by three staff members in 2012.

coordination between regional M&E related networks and global partners. Finally, CLEAR provided a lot of visibility, recognition and learning focus on local innovations on M&E in the region.

In 2015, CLEAR LAC launched the first Evaluation Week in Mexico, which was expanded to Latin America in 2017. In 2018, more than 100 institutions hosted events for the Evaluation Week, which entailed an audience of over 10,000 people in 14 Latin American countries. In 2019, this initiative went global, with the participation of subnational governments, NGO's, networks and national government agencies from 37 countries worldwide. Over time, the careful documentation of every event and participating institution have provided a good outlook of the regional agenda on M&E and further promoted communication, learning and cross-fertilisation (Table 13.1). In comparative perspective, these

**Table 13.1** CLEAR strategy 2013–2018 (Universalia Management Group, 2014)

---

**CLEAR Highest-Level Outcome:** To contribute to stakeholders in the target regions using evidence in making decisions for improved development results

**CLEAR Higher-Level Outcome:** Strengthened context-specific M&E systems and practices

**CLEAR's Immediate Intended Outcomes:**

- Improved enabling environments and strengthened demand for M&E
  - Critical mass of professional expertise
  - Innovations in M&E
- 

results are encouraging. Naturally, they do not only reflect the individual effort or merit of the CLEAR Centre, they are clearly a function of the enabling environment in Latin America and the inclusiveness and collaborative attitude of pre-existing networks and partners in the development cooperation and the M&E field.

First of all, CLEAR LAC was beneficiary of a host country effect: at that time, Mexico was not only proactively promoting this agenda domestically, but it was increasingly and purposefully becoming an international reference for other countries (the case of CONEVAL).<sup>18</sup> In fact, the Mexican government selected M&E as a national priority to guide international cooperation, mainly South-South. In the previous decades, under

<sup>18</sup> National Council for the Evaluation of the Social Development Policy.

the aegis of the so-called third wave of democratisation (Huntington, 1991) and structural reform efforts many Latin American countries had advanced accountability mechanisms and implemented better and more credible government information systems (Pérez & Maldonado, 2015). Governments in the region experienced increased electoral competition and pluralism (pressures from below) and increasing participation of the multilateral financial institutions in the national policy dialogue, arguably as a result of the Washington Consensus (pressures from abroad). It is undeniable that during this period, the region as a whole advanced in the adoption of results-based management tools including monitoring, evaluation, as well as performance-based budget, performance audits and governmental metrics of impact and development progress, with varying degrees of success and institutionalisation.<sup>19</sup>

In my view, the trajectory of CLEAR LAC suggests that the global theory of change was particularly suitable for the Latin American context: a high level of governmental demand for M&E, the nature and scope of civil society networks, dynamic regional institutional spaces, as well as a closely intertwined network of M&E professionals working in the region were all features that strongly favoured the capacity-building strategy championed by CLEAR. These actors had been anticipating a rising demand in evaluation for many years and had already made important contributions to the regional policy dialogue. With CLEAR, their efforts gained visibility and diffusion, and the Centre helped mobilise informal issue networks that enabled useful information exchange and a stronger understanding of the regional outlook of M&E among stakeholders. In addition to core courses of M&E (evaluation methodology, development of indicators, planning for result, PBM, evaluation management, communication of results, etc.) for different audiences and levels of expertise, the Centre has provided a stable source of knowledge products in Spanish as a public good and has launched an online platform (APRENDER) that seeks to include exhaustive information on the suppliers of ECB. In the context of this enabling environment, CLEAR was the only actor with incentives and resources to strengthen and disseminate these efforts on a regional basis and identifying cross-cutting issues for the promotion of the evaluation culture.

<sup>19</sup> For a detailed account, see Ospina, S. N., Cunill-Grau, & Maldonado, C. (2021). *Enhancing accountability through results-oriented monitoring and evaluation systems*. Handbook of Latin American public administration, Yorkshire: Emerald Group.

Furthermore, CLEAR LAC can reasonably be credited with a contribution to institutionalisation and capacity development for evaluation in the region. Despite relevant country-level differences, during this period, the Centre provided, directly and indirectly, a vast array of services to virtually every country in the region. With the benefit of hindsight, it is clear that the Centre provided a unique combination of services and linkages thanks to the convening power of an academic institution, on the one hand, and the flexibility with which it approached governments, legislators and non-governmental actors alike, on the other. The combination of institutional anchoring and resources and incentives for proactive mobilisation of stakeholders turned out to work relatively well.

### CONTRIBUTIONS TO ECB AND INSTITUTIONALISATION

In Latin America, the demand for evaluation and results-oriented management tools had been part of the agenda of State modernisation for at least two decades. At the same time, rising demands for accountability and new forms of social mobilisation and political participation had emerged as a result of democratic competition. M&E, however, had remained largely a governmental tool to control budgets and promote efficiency, rather than a broader instrument for public accountability and public dialogue (Ospina et al., in press). This democratic deficit in the system entailed a structural weakness of the relatively well-developed institutional frameworks for evaluation. These institutional frameworks and legal structures were anchored in a high-turnover bureaucracy, in the absence of a strong civil service tradition. This entailed—*de facto*—a large scope for manoeuvre within the institutional framework, the loss of institutional routines and legacies, organisational know-how and high probability of discontinuity in the evaluation function via personnel turnover.

By incorporating a multidimensional and more consciously context-sensitive and dynamic notion of capacity building, CLEAR LAC contributed to institutionalisation by promoting knowledge exchange for institution-building among key governmental actors (line ministries, legislators, planning ministries), and simultaneously supporting linkages and issue networks beyond the governmental sphere that could accompany these efforts and empower an informed demand for results-based accountability and democratic dialogue (NGO's, issue networks, watchdog organisations, professional associations, etc.)

As Caroline Heider (2011) suggests, the concept of evaluation capacity development (ECD) evolved from a focus on *just training* to the recognition that effective capacity building entailed the need to simultaneously intervene at least three interrelated levels: (1) individual; (2) organisational and institutional and (3) the enabling environment. So conceived, capacity building also entails the need to create synergies between the evaluation profession and expertise in capacity development, and placed further emphasis on the need to promote sustainability of efforts and practice beyond individuals and institutional frameworks. In a similar fashion, Baizermann et al. define capacity building as “intentional work to continuously create and sustain overall organisational processes that make quality evaluation and its uses routine” (2002, p. 109). By providing diversified services (beyond training) and exploring both formal and informal linkages with regional and non-governmental actors, CLEAR LACs activities have been clearly aligned with known mechanisms for capacity building:

ECB involves the design and implementation of **teaching and learning strategies** to help individuals, groups, and organisations learn about what constitutes **effective, useful, and professional evaluation practice**. The ultimate goal of ECB is sustainable evaluation practices –where members continuously ask questions that matter; collect, analyse, and interpret data; and **use evaluation findings for decision-making and action**. For evaluation practices to be sustainable, organisation members must be provided **leadership support, incentives, resources, and opportunities to transfer their learning about evaluation** on their everyday work. Sustainable evaluation practice also requires the development of systems, processes, policies and plans that help embed evaluation work into the way the organisation accomplished its strategic mission and goals. (Preskill & Boyle, 2008b, p. 49)<sup>20</sup>

In addition to training services, the Centre promoted the culture of evaluation by generating knowledge products in Spanish; sponsoring awards, internships and recognition for innovation in evaluation, leadership opportunities; showcasing good practice and the benefits of knowledge exchange; actively promoting the use of evidence for decision-making, enabling dialogue and empowering emerging communities of evaluation,

<sup>20</sup> See also Table 13.1.

as well as providing citizen watchdog organisations with the conceptual and methodological baggage required to mobilise results-oriented demands for accountability.

This diversified and more horizontal approach to ECB had long been advocated in the development world.<sup>21</sup> UNDP, for instance, anchored its NEC initiative in a very similar fashion:

The aim was to provide a forum to discuss issues that face programme country partners, to deepen their understanding of evaluation as a powerful tool for public accountability, to learn from solutions in other developing countries and, possibly, to identify common strategies to establish evaluation systems that are relevant and have a sound political and institutional basis. It also aimed to provide a platform for sharing experiences in strengthening the institutional capacities and/or enabling conditions for evaluation of public policies, programmes and institutions. (UNDP, 2011, p. 2)

This shared view among ECB providers and M&E stakeholders responded to the known limitations of foreign-driven reforms and government-led evaluation, and the adoption of a more learning-oriented and context-based ECD (Table 13.2). Ownership, a focus on learning and the aspiration of cross-fertilisation within the global South best capture this new ethos, which entailed the need for diversified strategies and more reliance on soft power (i.e. informal networks and communities of practice).

At the same time, this modality of ECB also requires a much larger scale and scope to be able to visibly affect institutionalisation and democratic decision-making in the desired direction. If we define institutionalisation as “the incidence and permanent anchoring of formal and informal rules, structures and processes that guide and stabilise the behaviour of individuals and groups, thereby regulating societal behaviour” (Stockmann et al., 2017), one would need to acknowledge that despite the fact that M&E systems in Latin America are legally formalised for the most part, this does not amount to permanent anchoring and/or consolidation. Common challenges are related to deficits in human capital (which leads to low quality of evaluations), the underutilisation of evaluation for

<sup>21</sup> Another example of this is the CoPLACGpRD, a community of practice of experts in Results-Based Management for Development from Latin America and the Caribbean.



**Table 13.2** Strategies for ECD (Authors’ elaboration based on Preskill & Boyle, 2008a, p. 157)

<i>10 ECD Strategies</i>	<i>CLEAR’s context-specific regional strategy</i>
Involvement	• Permanent Seminar and Outreach
Training	• Diploma on Public Policy and Evaluation
Technical Assistance	• Evaluations, indicators development, usability of recommendations, etc
Written Materials	• Evaluation library in Spanish (Anthology Series and over REDLACME)
Communities of Practice	• M&E regional and national networks, emergence of National Academy of Evaluation in Mexico, VOPES
Appreciative Inquiry	• <i>n/a</i>
Technology	• Dissemination of the role of ITC in M&E
Internships	• CLEAR summer internship programme for Latin Americans
Mentoring	• M&E graduate research award for Latin Americans
Meetings	• National Evaluation Capacities, CLAD, REDLACME, IDEAS

decision-making, lack of citizen interest and legislators’ involvement, as well as problems of fragmentation and lack of coordination of the entire management cycle.

In fact, as mentioned before, the basic tenets of the CLEAR initiative theory of change seem to have held during the last decade. Background conditions seemed to provide fertile ground for the promotion of the evaluation agenda following the ‘CLEAR way’, which basically entails that CLEAR is a good pro-cyclical intervention in the promotion of M&E: it works well when things are going well. However, can CLEAR be expected to serve as an enabler and convener for the promotion of evaluation when policy discourse is going against the current of results-oriented accountability? Can ECB thrive if past trends towards high demand for evaluation and more institutionalisation of these tools are reversed?

The sustainability model of CLEAR LAC is anchored in a cross-subsidisation model and management strategy that is aimed at increasing the weight of fee-based services and interventions to be able to subsidise the provision of public goods. As long as regional demand for evaluation training and services remained high, the long-term sustainability of CLEAR seemed likely, considering the good reputation of CLEAR trainings and widespread recognition of its knowledge products (i.e.

Anthology Series). However, donors' expectations that Centres could become financially self-sustainable by the end of phase II could only be potentially met if increasing trends in demand prevailed and Centres became pivotal actors in mobilising new demands for ECB. In a context of contracting government demand, however, the risks for sustainability are evident and the need to subsidise dissemination and advocacy efforts becomes more acute. In their current form, Centres do not have enough resources and the leverage needed to defend results-oriented accountability from government policy. Among other factors, Centres have insufficient funding to accommodate within-region heterogeneity and are often forced to adopt cream-skimming strategies rather than invest in more challenging endeavours.

The emergence of populist governments in many countries,<sup>22</sup> as well as post-truth leadership styles in the region, and a general anti-expert sentiment in public opinion clearly posit a great threat to the institutionalisation of evaluation. Widespread mistrust in expert policy advice, statistical information, scientific evidence and a rejection of technical expertise with anti-elite arguments are all common ingredients in the public debate of many countries in Latin America—and beyond. For this reason, prospects for the long-term sustainability of CLEAR and its continued contribution to institutionalisation are a function of the political economy of the region and the willingness of donors and partners to respond to a changing environment in the region, and the need to further invest in these types of initiatives, even against the current.

### LIST OF ABBREVIATIONS

CLAD	Centro Latinoamericano de Administración para el Desarrollo
CLEAR LAC	Center for Learning on Evaluation and Results for Latin America and the Caribbean (Mexico)
CIPPEC	Centro de Implementación de Políticas Públicas para la Equidad y el Crecimiento (Argentina)
Coneval	Consejo Nacional para la Evaluación de la Política de Desarrollo Social
DFID	Department for International Development (U.K.)

<sup>22</sup> Mexico, Brazil, Ecuador, Bolivia and El Salvador, to name a few.

ESAP	Escuela Superior de Administración Pública (Colombia)
GRADE	Grupo de Análisis para el Desarrollo (Peru)
FOCEVAL	Fomento de Capacidades en Evaluación
IEG	Independent Evaluation Group (World Bank)
ENAP	Escola Nacional de Administração Pública (Brasil)
IDEAS	International Development Evaluation Association
M&E	Monitoring and evaluation
NEC	National Evaluation Capacities
REDLACME	Red Latinoamericana y del Caribe de Evaluación y Monitoreo
SIDA	Swedish International Cooperation Agency

## REFERENCES

- Baizermann, M., Compton, D. W., & Stockhill, S. (2002). New directions for ECB. *New Directions for Evaluation*, 93, 109–119.
- Heider, C. (2011). Conceptual framework for developing evaluation capacities: Building on good practice in evaluation and capacity development. In R. C. Rist, M. H. Boily, & F. Martin (Eds.), *Influencing change: Building evaluation capacity to strengthen governance* (pp. 85–107). The World Bank Group.
- Huntington, S. P. (1991). *The third wave: Democratisation in the late 20th century*. University of Oklahoma Press.
- OECD. (2012). Aid effectiveness 2011. *Progress in Implementing the Paris Declaration, Better Aid, OECD Publishing, Paris*, 2018. <https://doi.org/10.1787/9789264125780-en>. Accessed on 3 October.
- Ospina, S., Cunill-Gra, N., & Maldonado, C. (2021). *Enhancing accountability through results-oriented monitoring and evaluation systems*. Handbook of Latin American Public Administration. Yorkshire: Emerald Group.
- Pérez, G., & Maldonado, C. (2015). *Panorama de los Sistemas Nacionales de Evaluación y Monitoreo en América Latina*. CIDE.
- Pérez, Y., Maldonado, C., & Faustino, D. G. (2015). El Sistema de seguimiento y evaluación de programas federales en México: Retos para su consolidación. In Y. Pérez, & C. Maldonado (Eds.), *Panorama de los sistemas nacionales de monitoreo y evaluación en América Latina* (pp. 273–311). México: CIDE.
- Preskill, H., & Boyle, S. (2008a). Insights into evaluation capacity building: Motivations, strategies, outcomes, and lessons learned. *The Canadian Journal of Program Evaluation*, 23(3), 147–174.

- Preskill, H., & Boyle, S. (2008b). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, 29(4), 443–459.
- Stockmann, R., Meyer, W., & Taube, L. (2017). *Globo de Evaluación. Concepto General*. Saarbruecken: Center for Evaluation [Internal and not published working paper].
- UNDP, (2011). Proceedings from the International Conference on National Evaluation Capacities (NEC 2009). 15–17December 2009. Kingdom of Morocco.
- Universalia Management Group. (2014). *CLEAR mid term evaluation report*.



# Contributions of the Inter-American Development Bank Group (IDB Group) to the Institutionalisation of Evaluation in Latin America and the Caribbean

*Ana Maria Linares, Anna Funaro Mortara,  
and Melanie Putic*

## INTRODUCTION: EVALUATION AT THE IDB GROUP

In the 1970's multilateral organisations sparked the practice of evaluation in the international development field in LAC as an *ex ante* procedure during project design and early implementation (Cunill-Grau & Ospina, 2012).

---

The opinions expressed in this article represent those of the authors and not necessarily those of IDB Group.

---

### *Present Address:*

A. M. Linares (✉) · M. Putic  
Office of Evaluation and Oversight (OVE), Inter-American Development Bank,  
Washington, DC, NW, USA  
e-mail: [anamariali@iadb.org](mailto:anamariali@iadb.org)

Throughout the 1980s and 1990s, the focus of evaluation began to broaden to include *ex post* analyses. The discipline then experienced increasing efforts towards formalisation, capacity development initiatives, and the drafting of early evaluation technical frameworks. The new millennium saw the practice of evaluation expand beyond multilateral institutions to governments, applied to the analysis of public policies and national development plans and increased participation of civil society in evaluation activities (Burdescu et al., 2005; Neirotti, 2012).

Evaluation has been an integral part of IDB Group's<sup>1</sup> work since the late 1990s. It has evolved to become the system it is today, in which evaluation is a shared responsibility between IDB Group management and the Independent Office of Evaluation and Oversight (OVE).<sup>2</sup> OVE was created in 1999 as an independent evaluation office of IDB, separate from IDB management's internal office of evaluation existing at the time. As per the Organisation for Economic Co-operation and Development's (OECD) definition, independent evaluations are "carried out by entities and persons that are organisationally and behaviourally independent from those responsible for the design and implementation of the intervention" (2002, p. 24). OVE reports directly to the Board of Executive Directors and is responsible for evaluating the development effectiveness and performance of IDB Group's work in LAC, undertaking independent and systematic evaluations of IDB Group's strategies, policies, programmes, operations, activities, and systems. In 2016, following the consolidation of IDB Group's private sector windows, OVE also became the independent evaluation office of IDB Invest, reporting to its Board of Executive Directors.

On IDB Group management's side, IDB's internal evaluation function has been the responsibility of the Strategic Planning and Development

---

M. Putic  
e-mail: [melanick@iadb.org](mailto:melanick@iadb.org)

A. F. Mortara  
Washington, DC, NW, USA

<sup>1</sup> The IDB Group comprises the Inter-American Development Bank, responsible for public sector lending and assistance, and the Inter American Investment Corporation (IDB Invest), responsible for private sector lending.

<sup>2</sup> <https://www.iadb.org/en/ove/home>.

Effectiveness Department (SPD) since 2007. Following the realignment of the IDB in 2007, SPD replaced the existing evaluation office and expanded its scope. SPD focuses on project evaluations and is responsible for the self-evaluation component of the IDB's evaluation system. As defined by the OECD, self-evaluations are conducted "by those who are entrusted with the design and delivery of a development intervention" (2002, p. 31). Since 2016, the Development Effectiveness Division (DVE) has served the same function for IDB Invest. At the core of self-evaluation systems across the IDB Group are evaluations of development interventions (e.g., public sector operations, private sector operations, and advisory services), which are embedded in the project cycle. Self-evaluations at project completion build on ex ante evaluability assessments and implementation monitoring, which are an integral part of IDB Group's project preparation and portfolio management processes. Both IDB<sup>3</sup> and IDB Invest<sup>4</sup> self-evaluate the performance and results of their projects. OVE reviews and validates these self-evaluations, assigning final performance ratings to each project, which are then used for IDB Group's corporate reporting. Finally, operational units and staff within IDB Group play an important role in fostering evaluation in the LAC region by including the evaluation of interventions in IDB Group-financed loans and technical assistance projects.

In 2019, the IDB Group brought together various aspects of its evaluation system in one reference document, the Evaluation Policy Framework, approved by the Boards of Directors of both IDB and IDB Invest. According to this framework, the objectives of the IDB Group's evaluation system are to improve the "operational performance, strengthen institutional learning, and achieve better results from evaluation working together with development partners" (IDB, 2019a, p. 2). As such, the evaluation system is expected to help ensure that IDB and IDB Invest are

<sup>3</sup> Self-evaluation at the IDB was launched with the Development Effectiveness Framework (DEF) in 2008. The DEF is reviewed and updated periodically, with the most recent update in 2018 (IDB, 2018).

<sup>4</sup> For IDB Invest, the more systematic tracking of development results dates back to 2001 when the IIC established a project evaluation system and function in 1999, contracted the OVE to support this effort in 2000, and in 2001 started to self-evaluate its operations following the Good Practice Standards for Private Sector Operations issued by the MDBs' Evaluation Cooperation Group (ECG, 2006). In 2016, OVE became the office of evaluation and oversight of IDB Invest following the consolidation of IDB Group's private sector operations into IDB Invest.

effective in providing development results within the region. To achieve this goal, the IDB Group needs to continuously learn and adapt, based on evidence of what does and does not work. Evaluation supports such adaptive learning, fosters accountability for achieving results and contributes to institutional transparency.

## IDB GROUP'S SUPPORT TO INSTITUTIONALISING EVALUATION IN LAC

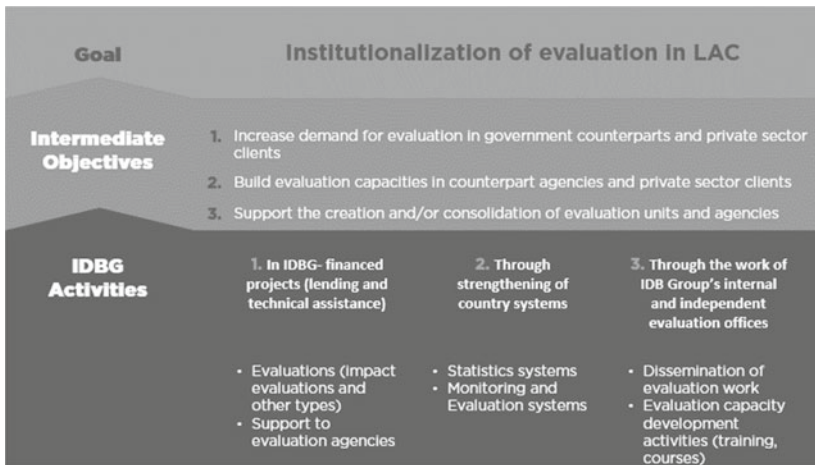
### *Conceptual Framework*

The IDB Group is the leading source of development financing for LAC, providing financial and technical support to reduce poverty and inequality in the region. The purpose of the IDB, as stated in the 1959 Agreement Establishing the Inter-American Development Bank, is to “contribute to the acceleration of the process of economic and social development of the regional developing member countries, individually and collectively.” (IDB, 1959, Art. 1, Sec. 1, p. 5).

The purpose of IDB Invest, as stated in the Agreement Establishing the Inter-American Investment Corporation, is “to promote the economic development of its regional developing member countries by encouraging the establishment, expansion, and modernisation of private enterprises.” (Inter-American Investment Corporation, 1986, Art. 1, Sec. 1, p. 1). Through its work, IDB Group supports the institutionalisation of evaluation in LAC in at least three ways. First, through learning by doing, as part of IDB Group-financed projects. This support can take the form of evaluations of certain aspects of policy or administrative reforms, of pilot initiatives, or of existing interventions. It can also take the form of strengthening and sometimes creating evaluation units in government counterparts and private sector clients to monitor project implementation and evaluate project results at completion. Second, through its strategy on Country Systems, IDB Group supports the strengthening of agencies responsible for generating accurate, reliable and timely statistics, and of agencies in charge of M&E systems, both of which are key building blocks for evaluation to take root permanently (IDB, 2009; subsequently updated in IDB, 2019b). Third, through the work of its evaluation offices, both internal and the independent evaluation office, IDB Group generates evaluation work that is widely disseminated throughout LAC to promote learning. In addition, the evaluation offices conduct various



activities to generate evaluation capacity in LAC counterparts and clients, such as training and courses. Together, these three streams of work contribute to (1) increasing the demand for evaluation in government agencies and in private sector clients, (2) building evaluation capacities within counterparts, and (3) supporting the creation and/or consolidation of evaluation units, all of which advance the institutionalisation of evaluation in LAC countries. Figure 14.1 illustrates this theory of change.



**Fig. 14.1** Theory of change underlying IDB group's work to support institutionalisation of evaluation in LAC (OVE)

### *Support to the Institutionalisation of Evaluation Through IDB Group-Financed Projects*

Many investment projects and technical assistance operations, particularly with public sector agencies but also with private clients, include the evaluation of specific interventions or certain aspects of policy or institutional changes, pilot interventions and established programmes that IDB Group finances. These evaluations cover the gamut of evaluation types, from process evaluations to impact evaluations and other types in between, that are an integral part of some operations. They generate evaluation capacity in IDB Group's counterparts, as they focus directly on

borrower and beneficiary activities and are implemented by the borrower or the beneficiary, either directly or through hiring of consultants. IDB Group's specialists provide support from the initial phases of evaluations (drafting TORs for consultants and assisting in the design of the evaluation) and supervise their implementation until completion. One example is the *Programme for Education, Health and Food* (PROGRESA),<sup>5</sup> Mexico's cash transfer programme, which IDB funded and closely followed through a series of evaluations. These evaluations have provided evidence of the programme's positive results and helped make adjustments to it over time. They also generated capacity within the government to adopt the practice of continuous evaluation of the programme as it evolved, which continues to this day.

IDB Group does not maintain a central database or official registry of all the evaluations financed through its lending and technical assistance operations. Nonetheless, in 2017, OVE conducted an evaluation of one such type, IDB's impact evaluations, and assessed their production and use in public sector operations (IDB, 2017). OVE found that between 2006 and 2016, more than 400 projects had proposed to conduct at least one impact evaluation. IDB had budgeted approximately US\$200 million to finance those evaluations, although about 25% of that was not used, as evaluations were cancelled due to political challenges or implementation and design issues. OVE also found that most government counterparts that participated in an IDB Group-financed impact evaluation were satisfied, either because the evaluation helped ensure the sustainability or continuity of the project, or because it created capacity for the institution to continue conducting this type of evaluation. OVE's analysis showed that the data requirements of impact evaluations financed by IDB led to improvements in the M&E systems of government counterparts and in their reporting activities. In fact, two-thirds of the interviewees agreed that after having conducted an IDB-financed impact evaluation, they were better equipped to conduct other impact evaluations in the future, thanks to the training provided by the Bank.

Another way IDB Group-financed operations support the institutionalisation of evaluation is through creation and/or consolidation of evaluation units in public sector counterpart agencies. Many IDB Group projects require the establishment or strengthening of a project executing

<sup>5</sup> <https://www.gob.mx/cms/uploads/attachment/file/79893/1998.pdf>.

unit by incorporating one or more M&E specialist. These M&E specialists are responsible for implementing the project's M&E plan, gathering the baseline and monitoring data during project implementation to feed the final evaluation that will assess the overall effectiveness of the intervention. Sometimes these M&E specialists remain in the executing agency after project completion and are absorbed as part of an M&E unit within the agency, thus directly contributing to the institutionalisation of evaluation. A case in point is the evaluation unit of the Programme for Investment in Agricultural Services of the Ministry of Agriculture in Argentina. Through various loans from by the IDB<sup>6</sup> and the World Bank, since the early 2000s, this evaluation unit was created and subsequently fully absorbed by the Ministry. Today it undertakes a wide range of evaluations on its own.<sup>7</sup>

### *Support to Institutionalisation of Evaluation Through Strengthening of Country Systems*

In 2009 IDB approved a Strategy for Strengthening and Use of Country Systems (GN-2538) that defined how the IDB would work towards strengthening its clients' national systems and increase their use in Bank-financed operations. The Strategy represented IDB's response to the Paris Declaration on Aid Effectiveness (2005) and the Accra Agenda for Action (2008), which emphasised that success in development depended in large measure on governments' capacities for implementing their policies and managing their public resources through their own institutions and systems. After ten years of implementation, IDB approved an Update of the Strategy in 2019 (GN-2538–31). The term *country systems* refers to all national norms, regulations, procedures, and structures for the management of the public sector. Country systems include *fiduciary* and *development effectiveness* systems. Fiduciary systems are those for procurement and financial management and, to date, have received IDB's priority attention. Development effectiveness systems are those for strategic planning, M&E, statistics, and environmental and social safeguards. Support for M&E and for statistics systems is a foundation for institutionalising

<sup>6</sup> <https://www.iadb.org/en/project/AR-L1030>.

<sup>7</sup> [http://www.prosap.gov.ar/se\\_descripcion.aspx#](http://www.prosap.gov.ar/se_descripcion.aspx#).

evaluation in LAC countries. M&E systems provide the indicators, monitoring tools, reporting mechanism, and results assessments that should feed evidence-based public policies. For their part, statistical systems provide the data needed for M&E to work.

IDB Group actions to strengthen country systems are typically financed with non-reimbursable technical assistance resources, though they can also be financed through lending as part of institutional capacity-building activities in loan projects. Most of IDB's work on M&E systems has focused on supporting monitoring systems. M&E systems at the IDB are those used to quantify the goods and services provided by the public sector, determine their quality and targeting, and measure the effect of the results obtained. Updated of the Strategy for Strengthening and Use of Country Systems (GN-2538–31). Support to strengthening statistical systems of countries in LAC has mostly taken the form of specific investment loans with interventions directed at the institutional structure of national statistics offices; data production; the use of information; and innovation. Examples of these interventions include (1) the Programme in Argentina to Strengthen the Statistical Capacity of the National Statistics and Census Institute<sup>8</sup> where IDB financing was used to upgrade the Institute's information technology infrastructure, and to update the sociodemographic statistical database, among other uses; and (2) the Programme in El Salvador to Support the National Information and Statistics System,<sup>9</sup> through which the national statistics database was updated, among other things.

### *Support to Institutionalisation of Evaluation Through IDB Group's Evaluation Offices*

The roles and responsibilities of IDB and IDB Invest management concerning self-evaluation, and of OVE for independent evaluation, explicitly include the need to support evaluation capacity development in LAC. Although the main responsibilities of IDB<sup>10</sup> and IDB Invest<sup>11</sup>

<sup>8</sup> <https://www.iadb.org/en/project/AR-L1266>.

<sup>9</sup> Operation ES-L1128: Support to the National Statistics and Information System and National Planning System (El Salvador).

<sup>10</sup> Self-evaluation at IDB is the responsibility of the Strategic Planning and Development Effectiveness Department or SPD.

<sup>11</sup> Self-evaluation at IDB Invest is the responsibility of its DVE.

within the evaluation system involve ensuring effective self-evaluation, they also include supporting borrowers and clients in their evaluation activities through joint evaluations and technical support, and preparing knowledge products based on M&E findings and disseminating them within and outside the IDB Group to promote learning on most effective solutions and approaches to development. As for OVE, in addition to its main responsibilities regarding undertaking independent and systematic evaluations of the IDB Group's performance and the development effectiveness of its work, it has a mandate to encourage and support local counterparts in the LAC region in building effective M&E systems and capacity, and to support a learning environment based on cooperation, information-sharing and constructive feedback.

#### *Support Through IDB Group's Internal Evaluation*

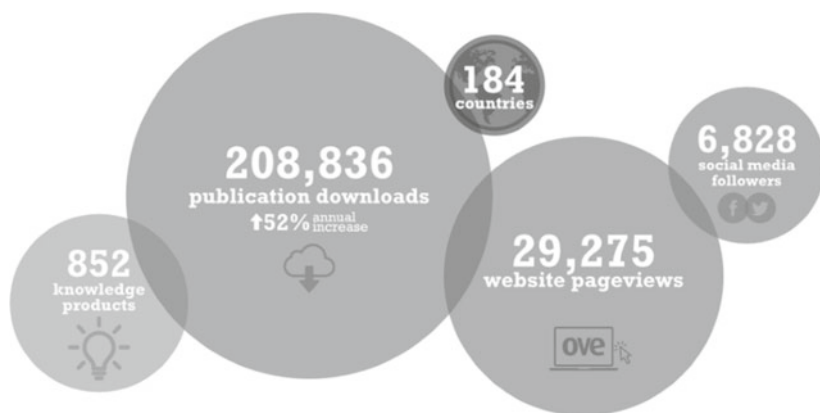
IDB's commitment to promote learning in general, and more specifically evaluation learning, has prompted the creation of a variety of evaluation tools and resources that are made available to external audiences. For example, in 2018 SPD created the Evaluation Hub website<sup>12</sup> as a repository of different evaluation tools, guidelines and other resources on impact and economic evaluation available to users for consultation and use. Based on the premise that the generation and use of rigorous evidence is key to the effectiveness of development interventions, the Hub was created to place evaluation sources, resources, and new developments at the fingertips of practitioners and other interested parties. Another example includes training sessions conducted by SPD (Impact Evaluation Design Week)<sup>13</sup> for IDB staff involved in the design of a new project, to which government counterparts are invited. With the support of SPD specialists, project teams learn about impact evaluation, and those interested in incorporating it into the project's design receive hands-on support. Finally, IDB's project completion reports, which are made available to the public, provide another source for learning and strengthening evaluation capacities in the LAC region.

<sup>12</sup> <https://www.iadb.org/en/evaluationhub>.

<sup>13</sup> <https://www.iadb.org/en/topics-effectiveness-improving-lives/impact-evaluation-design-week-jul-aug-2018>.

### *Support Through IDB Group's Independent Evaluation*

OVE's outreach activities and dissemination of evaluations and their findings beyond the IDB Group are geared towards supporting evaluation capacity development in the LAC region. OVE seeks to stimulate learning through its outreach and dissemination activities, using a variety of channels to reach government counterparts, private sector clients, academic institutions, and the public at large. OVE makes the content of its evaluations available to the public online through its website, manages a subscriber base, sends newsletters to targeted audiences, maintains an active presence on social media, distributes brochures and other outreach materials, sponsors events, and participates in a variety of workshops and conferences. In 2019, for example, OVE registered an increase in the number of downloads of evaluation reports and knowledge products. OVE's total report downloads rose to 208,836; 50% more than in 2018 (Fig. 14.2). Downloads were made from all over the world, but particularly from countries in the region—mainly Colombia, Mexico, Peru, Argentina, Ecuador, Bolivia, and Chile (IDB, 2020).



**Fig. 14.2** Outreach of the office of evaluation and oversight (OVE, IDB Knowledge Analytics, 2019)

In addition, OVE has sought to leverage its efforts towards creating evaluation capacity in LAC through its support to the Centers for Learning on Evaluation and Results, known as the CLEAR<sup>14</sup> initiative. The CLEAR Initiative, established in 2010, is aimed at helping governments use and apply the principles of M&E to decision-making processes (CLEAR, 2012). The CLEAR initiative has six regional centres around the world, of which OVE supports the two that are based in LAC. The Centre for Research and Teaching in Economics (CIDE) in Mexico was the first Spanish-speaking CLEAR centre for LAC, and the Getulio Vargas Foundation in São Paulo (Brazil) is the Portuguese-speaking CLEAR Center. As a global programme bringing together academic institutions and donor partners, the CLEAR centres in Mexico and Brazil focus on strengthening local capacities in M&E through training, technical assistance, and research, generating evaluation expertise and disseminating evaluation knowledge for LAC (CLEAR, 2018; Pérez Yarahuán et al., 2019). In 2019 for example, the work with the CLEAR Centers from Mexico and Brazil contributed to strengthening evaluation capacity in the LAC region. CLEAR/CIDE developed an assessment methodology for M&E capacities for social development agencies that is being implemented in 19 countries in LAC. Furthermore, CLEAR/Getulio Vargas Foundation developed a methodology to implement M&E systems at the state level.<sup>15</sup> Worldwide, in June of 2019, the CLEAR initiative launched the first *Global Evaluation Week*, which organised knowledge dissemination events about M&E at the global scale, fostering the exchange of knowledge on the subject. OVE takes part in these events as a regional sponsor to encourage the participation of policy makers and civil society from the LAC region.

## CONCLUSION

The LAC region is experiencing a promising trend towards a broader evaluation reach. As shown in this article, IDB Group's various activities contribute in different ways to the institutionalisation of evaluation in the LAC region. These efforts notwithstanding, consistently measuring and

<sup>14</sup> <https://www.theclearinitiative.org/>.

<sup>15</sup> Work with Espírito Santo State was concluded in 2019, but the work with Minas Gerais State is not yet finished. From these experiences, an intervention strategy for other Brazilian states will be developed.

assessing results of interventions and programmes and making effective use of evaluation findings remains a challenge. In addition, there is room to enhance IDB Group's contribution in coordination with other key players in the region, such as academic institutions, evaluation agencies and multilateral development agencies, through the continuous search for new ways to expand the scope and reach of evaluation capacity building activities as a building block for institutionalisation of evaluation in LAC.

Going forward, OVE is preparing a renewed strategy to expand the scope and reach of its activities in support of developing evaluation capacity in LAC, based on a three-prong approach: (1) support assessments to establish the degree to which evaluation has been institutionalised in all countries of LAC, as well as identify key stakeholders and potential new collaborators; (2) support the establishment of a regional network of evaluation agencies to promote learning and information sharing; and (3) pursue opportunities to create new partnerships with LAC development banks and other institutions, in parallel with its continued support to the CLEAR Mexico and Brazil centers. These activities will be closely monitored to ensure continuous learning.

### LIST OF ABBREVIATIONS

CIDE	Centre for Research and Teaching in Economics
CLEAR	Centers for Learning on Evaluation and Results
DEF	Development Effectiveness Framework
DVF	IDB Invest's Development Effectiveness Division
IDB (Group)	Inter-American Development Bank (Group)
IDB Invest	Inter-American Investment Corporation (formerly IIC)
LAC	Latin America and the Caribbean
M&E	Monitoring and evaluation
OECD	Organisation for Economic Co-operation and Development
OVE	Office of Evaluation and Oversight (of the IDB Group)
SPD	IDB's Office of Strategic Planning and Development Effectiveness



## REFERENCES

- Burdescu, R., del Villar, A., Mackay, K., Rojas, F., & Saavedra, J. (2005). *Institutionalizing monitoring & evaluation systems: Five experiences from Latin America*. World Bank.
- CLEAR. (2012, June). *State of the initiative report 2010*. <https://www.theclearinitiative.org/resources/clear-state-of-initiative-report-2010-2012>. Accessed on 21 November 2020.
- CLEAR. (2018). *Annual report: July 2017—June 2018*. [https://www.theclearinitiative.org/sites/default/files/2019-01/CLEAR\\_Annual\\_Report\\_2017-18.pdf](https://www.theclearinitiative.org/sites/default/files/2019-01/CLEAR_Annual_Report_2017-18.pdf). Accessed on 21 November 2020.
- Cunill-Grau, N., & Ospina, S. M. (2012). Performance measurement and evaluation systems: Institutionalizing accountability for governmental results in Latin America. In S. Kushner & E. Rotondo (Eds.), *Evaluation voices from Latin America* (pp. 77–91). American Evaluation Association.
- Evaluation Cooperation Group. (2006). *Good practice standards for the evaluation of private sector operations*. <https://www.ecgnet.org/document/good-practice-standards-evaluation-private-sector-operations-third-edition>. Accessed on 21 November 2020.
- IDB. (1959). *Agreement establishing the inter-American development bank*. <http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=781584>. Accessed on 21 November 2020.
- IDB. (2009). *Strategy for strengthening and use of country systems*. Document # GN-2538.
- IDB. (2017). IDB's impact evaluations: Production, use, and influence. *Office of Evaluation and Oversight*. Document #RE-512–1.
- IDB. (2018). *The updated development effectiveness framework for sovereign guaranteed operations (SGO)*. Document #GN-2489–10.
- IDB. (2019a). Evaluation policy framework—IDB group. *Office of Evaluation and Oversight*. Document #RE-538–5. <https://publications.iadb.org/en/evaluation-policy-framework-idb-group>. Accessed on 21 November 2020.
- IDB. (2019b). *Update of the strategy for strengthening and use of country systems*. Document #GN-2538–31.
- IDB. (2020). 2019 annual report. *Office of Evaluation and Oversight*. <https://publications.iadb.org/en/ove-annual-report-2019>. Accessed on 30 November 2020.
- Inter-American Investment Corporation. (1986). *Agreement establishing the inter-American investment corporation*. <https://idbinvest.org/en/download/4754>. Accessed on 5 November 2020.
- Neirotti, N. (2012). Evaluation in Latin America: Paradigms and practices. In S. Kushner & E. Rotondo (Eds.), *Evaluation voices from Latin America* (pp. 7–16). American Evaluation Association.

- OECD. (2002). *Glossary of key terms in evaluation and results based management*. <https://www.oecd.org/dac/evaluation/2754804.pdf>. Accessed on 21 November 2020.
- Pérez Yarahuán, G., Peña Habib, O., Herrera Galván, E., & Pinel Valerio, K. (2019). Regional institutional capacities in monitoring and evaluation: Latin American ministries of social development 2018 project presentation. *Center for Learning on Evaluation and Results, Latin America and the Caribbean* and *Centro de Investigación y Docencia Económicas*.



# The Independent Evaluation Group (IEG) of the World Bank Group: Influences on Evaluation Structures and Practices Globally and in the Americas

*Maurya West Meiers*

## INTRODUCTION

The topic of this paper involves the influence of the World Bank Group's Independent Evaluation Group (IEG) on the evaluation structures and practices in the Americas. But as the World Bank is an agency with a global mission, it is a difficult task to separate IEG's influence only to the Americas. In other cases, it is difficult to disconnect IEG's influence from that of other actors when working in partnership, such as with other agencies (e.g. the Inter-American Development Bank-IDB) or other World Bank departments. Additionally, the availability of existing research on IEG's influence *in the Americas* is not abundant. As such, this paper often presents a somewhat broader approach to also include IEG's influence to

---

M. West Meiers (✉)

World Bank Independent Evaluation Group, Washington, DC, USA

e-mail: [mwestmeiers@worldbank.org](mailto:mwestmeiers@worldbank.org)

© The Author(s), under exclusive license to Springer Nature  
Switzerland AG 2022

R. Stockmann et al. (eds.), *The Institutionalisation of Evaluation in the Americas*,

[https://doi.org/10.1007/978-3-030-81139-6\\_15](https://doi.org/10.1007/978-3-030-81139-6_15)

the field and practice of evaluation more globally, in particular in IEG's role as an independent evaluation body. The paper provides examples in the Americas where they are available and relevant.

To begin, an overview and history of the organisation are provided, including the objectives, vision and function not just of IEG, but also of the World Bank. The paper looks at the organisation in the region where it operates, with a focus on activities, organisation structure and so on. Again, this is a difficult task as the World Bank has 141 global office locations. But having the location of the World Bank Headquarters in the United States—in Washington, D.C.—and IEG's staff almost fully being based at Headquarters, IEG has no doubt been influenced by the North American setting. Finally, the paper examines the influence of its activities on the institutionalisation of structures and practices of evaluation in the Americas and globally.

This paper does not look at the evaluations that IEG conducts and the influence that these have had, which are beyond the scope of this paper's purpose.

## OVERVIEW OF IEG AND EVALUATION IN THE WORLD BANK GROUP: ESTABLISHMENT, STRUCTURE AND MANDATE, APPROACH, PURPOSE AND FUNCTIONS

### *About the World Bank Group*

The World Bank, founded in 1944 and headquartered in Washington, D.C., was established as the International Bank for Reconstruction and Development (IBRD) with the purpose of rebuilding countries following the destruction of World War II. The World Bank had an early emphasis on infrastructure developments, but over the years the World Bank's agenda has evolved. In the 1950s, the World Bank expanded its work beyond Europe to the developing world. Starting in the early 1960s, more emphasis was placed on working in the world's poorest countries, with a shift towards a primary purpose of poverty reduction. Over time, the World Bank Group entered other sectors such as agriculture, health and education. In the late 1990s, topics such as corruption, debt relief, disabilities, knowledge, the environment and gender came into greater focus.

Today, the World Bank Group is the largest development bank in the world, owned by 189 member countries, with offices in 141 countries and over 12,000 employees. It provides low-interest loans, zero to low-interest credits and grants to developing countries. Working with a range of partners—governments, other multilateral institutions, commercial banks, export credit agencies and private sector investors—the World Bank Group operates in areas such as education, health, public administration, infrastructure, financial and private sector development, agriculture, and environmental and natural resource management. The World Bank Group also works with middle-income and creditworthy low-income countries by providing loans, guarantees, risk management products and advisory services. The World Bank Group plays a role in coordinating responses to regional and global challenges. The agency’s current and two aligned goals are reducing poverty and building shared prosperity (World Bank, 2019a).

The World Bank Group is made up of five institutions. Two of these partner with governments—the IBRD and the International Development Association (IDA)—and three remaining institutions partner with the private sector—the International Finance Corporation (IFC), the Multilateral Investment Guarantee Agency (MIGA) and the International Centre for Settlement of Investment Disputes (ICSID). Together, IBRD and IDA form the World Bank, which provides financing, policy advice and technical assistance to governments of developing countries. IDA focuses on the world’s poorest countries, while IBRD assists middle-income and creditworthy poorer countries. IFC, MIGA and ICSID focus on strengthening the private sector in developing countries. Through these institutions, the World Bank Group<sup>1</sup> provides financing, technical assistance, political risk insurance and settlement of disputes to private enterprises, including financial institutions. This paper will mostly focus on the work with governments of the World Bank that is IBRD and IDA. Today, the World Bank operates day to day under the leadership and direction of a president, management and senior staff, and the vice presidents in charge of global practices (e.g. human development), cross-cutting solutions areas (e.g. gender), six regions and functions.

<sup>1</sup> The World Bank Group references all five institutions. The World Bank refers to IBRD and IDA. As the paper will not focus on IFC, MIGA or ICSID activities, the term ‘World Bank’ will be used hereafter.

### *The Establishment of the World Bank's Evaluation Function*

What is today known as IEG, as an independent body, began in 1973. But the World Bank began assessing projects in 1970 in the newly created Operations Evaluation Unit in the World Bank's Programming and Budgeting Department. In 1973, the unit became the Operations Evaluation Department (OED), making it the first independent evaluation function in an international financial institution.

The decision to begin evaluating projects was taken by a significant figure in the World Bank's history, President Robert McNamara (term 1968–1981), previously United States Secretary of Defense and Ford Motor Company President. McNamara's interest in evaluation had been informed by the increasingly routine use of evaluation in the 1950s and 1960s in the US government. At the time, programme assessments in a range of sectors, from education to criminal rehabilitation, had become more common in the US and Europe. While McNamara was US Secretary of Defense, the Planning, Programming, and Budgeting System (PPBS) was established within that department, with the goal of increasing efficiency and improving government operations. The budgeting process under PPBS emphasised objectives, planning and budgeting. The early efforts of the PPBS would eventually lead to the 'monitoring for results' movement (Morra Imas & Rist, 2009).

From its start as OED in 1973, until today, IEG as an evaluation body has reported to the Board of Executive Directors. After evaluation offices were established in two other private sector-focused bodies of the World Bank Group—in the IFC in 1984 and MIGA in 2002—the three evaluation functions were merged into the IEG in July 2006. As above, even though IFC and MIGA are part of IEG, this paper focuses mainly on the World Bank's work with governments.

### *IEG's Structure, Mandate and Independence*

IEG is headed by a term-limited Director General of Evaluation (DGE), who is approved by the Board of Directors, and operates under a formal mandate (World Bank, 2018). Through this mandate:

IEG is responsible for the assessment of *the relevance, efficacy, and efficiency* of World Bank Group operational policies, programmes and activities, and their contribution to development effectiveness.

IEG's evaluations aim to *enhance accountability and learning* to inform the formulation of new (and course corrections of ongoing) core work of the World Bank Group.

IEG *reports to the Board of Directors* and carries out its responsibilities and functions under the oversight of the Committee on Development Effectiveness (CODE).

An important aspect of this mandate is the issue of organisational structure, with the DGE and the IEG staff reporting to the Board of Directors, and not to World Bank Management (i.e. the World Bank President and any functions involving World Bank operations). IEG's work programme and budget are also approved by the Board of Directors. That being said, IEG does consult with World Bank Management in developing its work programme, seeking input on priority areas for evaluation, guidelines and so on.

Perhaps the most distinguishing aspect of IEG's structure and function is its status as an *independent* evaluation body. Four criteria are typically recognised as supporting independence of official evaluation and audit organisations: organisational independence; behavioural independence; avoidance of conflicts of interest; and protection from external influence (World Bank, 2003).

As specified under its current mandate, independence of evaluation in IEG is safeguarded through:

Structural independence: IEG reports directly to the Board of Directors

Institutional independence: The DGE is responsible for all functions, including managing the personnel and budget of IEG under the oversight of CODE.

IEG's work programme and budget are approved by the Boards.

IEG's evaluations are disclosed in line with its Access to Information policies and in accordance with the policies on Disclosure or Access to Information of the applicable Institution(s).

Managerial independence: IEG's DGE is appointed and removed by the Board, and IEG's staff is organisationally independent from the WBG operational entities.

The mandate further articulates arrangements for IEG and its staff on the avoidance of conflicts of interest.

### *IEG's Approach to Evaluation*

IEG's work 'centres' on evaluating development effectiveness. IEG uses a range of approaches to assess outcomes against stated objectives, benchmarks, standards and expectations, or assessing what might have happened in the absence of the project, programme or policy (counterfactual analysis). Across projects, IEG looks at the patterns of what works under what circumstances.

IEG's evaluation approach reflects, and is harmonised with, internationally accepted evaluation norms and principles, such as the quality standards for development evaluation of the Organisation for Economic Co-operation and Development (OECD), Development Assistance Committee (DAC), the good practice standards of the Evaluation Cooperation Group (ECG) and the norms and standards of the United Nations Evaluation Group (UNEG). In 2018 and 2019, IEG supported the effort on re-thinking the DAC evaluation criteria of relevance, effectiveness, efficiency, impact and sustainability, which have served in underpinning evaluations in many agencies.

In terms of quality assurance of its products and work, IEG relies on internal (IEG) and external peers, such as from academia, to review its evaluations. In 2016, a Methods Advisory Function was established, with the hiring of a Methods Advisor and supported by a small team, to promote internal knowledge sharing on evaluation design issues and methodological innovation. The establishment of this function formalised what had been a looser approach to reviews of methodological excellence. This Methods Advisory Function has since been adopted in the Asian Development Bank. Finally, IEG itself is periodically reviewed through an external panel, which is established and managed by the Board of Directors.

### *IEG's Purpose and Functions*

The purpose of IEG is to assess the performance of the institution's policies, projects and processes (accountability) and to learn what works in what context (learning). IEG carries out all independent evaluation work, mostly working on an ex post basis, and appraises the World Bank



Group's other evaluation systems and methods, including self-evaluation methodologies and results.

IEG's main evaluation work involves:

Evaluating whether the World Bank Group (WBG) is producing the expected results.

Incorporating evaluation assessments and findings into recommendations designed.

Reviewing the Institutions' operational self-evaluations.

Reporting regularly to the Boards on actions taken by World Bank Management in response to evaluation findings.

Identifying, disseminating and promoting the uptake of evaluation findings and lessons to reinforce learning and accountability within the WBG.

Encouraging and assisting member countries to build effective monitoring and evaluation partnerships, capacities and systems.

Another important feature of IEG's mandated work involves evaluation capacity development (ECD). Through its ECD efforts, IEG works closely with development partners and member countries in order to foster international evaluation harmonisation, to develop evaluation capacity in member countries and to encourage best practice in international development evaluation. More information about IEG's work on ECD is presented in the latter part of the paper.

### *IEG's Main Evaluation Products*

Through the years, IEG has a long history of producing standard evaluation products and approaches to evaluating development effectiveness. But these are adapted regularly, and new products are developed, to be responsive to changing World Bank Group operations and growth in its portfolio of products. Noting this summary does not do justice to the wider range of evaluation products that IEG produces each year, the core products can be grouped broadly as 'evaluations' and 'validations'. This 'evaluation' and 'validation' model has been adopted by other international agencies, especially in evaluation functions of multilateral and regional development banks.

To give a sense of scale and depth of IEG's work, in fiscal year 2019, the main evaluation activities involved five *major evaluations* (covering broad themes, sectors and corporate processes); two *meso evaluations* (which are 'just-in-time' evaluations that are focused on specific issues identified by the Board and Management); four *synthesis reports* or *reviews*; and seven *learning engagements*, and no *country programme evaluations*. IEG performed in-depth assessments of a sample of 43 projects through *project performance assessment reports*. IEG validated 443 project-level self-evaluations (including for IFC and MIGA projects) in the same year (World Bank, 2019b).

IEG's 'validations' involve reviewing the 'self-evaluations' that are completed by World Bank operational teams at the end of a World Bank-financed project. These are a significant part of IEG's work, given that all completed World Bank projects require self-evaluation.

The validation process for project-level self-evaluations (for IBRD and IDA projects) goes through multiple stages. All of this is done to assess the WBG's performance and identify lessons for improving World Bank Group operations at project level. As a first step during the design process, World Bank project task team leaders work with government counterparts to scope and design the project. As part of the documentation and approval process, each project requires (among many other things) a project development objective, a theory of change and an articulation of a results framework (with indicators, baseline data, targets, data sources, etc.). During the life of the project, various monitoring requirements must occur. The project team may also elect to commission or carry out its own evaluation activities. At the project's ending, all project teams must produce a project completion report (called an Implementation Completion and Results Report, or ICR). Once this ICR is produced, IEG undertakes a validation process for 100% of the World Bank's ICR. The IEG Implementation Completion and Results Report Review (ICRR) is an independent, desk-based, critical validation of the evidence, content, narrative and ratings included in the project team's ICR. Standard templates are used for both the ICR and ICRR. Based on the evidence provided in the ICR and an interview with the last task team leader, IEG arrives at its own ratings for the project using established criteria. Each ICR and ICRR are provided to both World Bank Management and the World Bank Board. The final ICR and ICRR ratings eventually become publicly available and become part of a publicly available data set that provides valuable trend data to track performance data

across a range of criteria. IEG's ICRRs serve as an independent validation of the results in the ICR and contribute to both learning and accountability.

### *Monitoring and Evaluation Functions Carried Out in the World Bank Group*

IEG is not the only body in the World Bank Group providing some form of monitoring and evaluation functions. In 2019, IEG co-developed the World Bank Group Evaluation Principles with counterparts representing World Bank Group Management and operations for IBRD, IDA, IFC and MIGA (World Bank, 2019c). These principles elaborate on the various functions and purposes of monitoring and evaluation (M&E) activities in the organisation.

Box 15.1 elaborates on how the World Bank Group distinguishes among three main evaluation modalities: independent evaluation, mandatory self-evaluation and demand-driven self-evaluation (World Bank, 2019c).

#### **Box 15.1: Types of Evaluations in the World Bank Group (World Bank, 2019c)**

**Independent evaluation.** A fully independent evaluation is carried out by entities that are structurally, functionally, and behaviourally independent from those responsible for the design and implementation of the intervention. In the World Bank Group, fully independent evaluations are conducted by IEG, which reports directly to the Boards. The emphasis is primarily on accountability and learning self-evaluation and impact evaluation at the level of the Boards or their Committee on Development Effectiveness (CODE). In addition, independent evaluations support learning at the levels of management and operations. The target audiences for independent evaluations are, therefore, the Boards, World Bank Group Management and staff, clients and development partners, as appropriate. In addition, to different degrees, independent evaluations also inform other actors such as (representatives of) beneficiaries and the general public.

**Self-evaluation.** Self-evaluations are conducted by operational staff or specific units within the management structures of the World Bank, IFC and MIGA and are therefore not fully independent of World Bank Group Management. They are usually closely linked to decision-making and

organisational learning processes within each institution. Self-evaluations are also conducted for purposes of accountability to World Bank Group Management and/or development partners/investors. The target audiences for self-evaluations are primarily operational units, management, clients and development partners. In addition, to different degrees self-evaluations also inform other actors such as (representatives of) beneficiaries and the general public. There are two broad types of self-evaluation in the World Bank Group.

**Mandatory self-evaluation.** At the core of the evaluation system across the World Bank Group are mandatory self-evaluations of specific lending operations, investments, guarantees, country programmes and advisory services. These evaluations are prepared by the responsible operational units and are embedded in the project and programme cycles. They are neither structurally nor functionally independent, but the principle of behavioural independence applies. Behavioural independence is further strengthened by IEG's review and validation (sometimes on a sample basis). Mandatory self-evaluations complement the implementation and monitoring arrangements that are embedded in each institution's project and portfolio management processes. Self-evaluation adheres to methods and guidance that are jointly accepted by World Bank Group Management and IEG, with predetermined concepts, formats and scope that are closely linked to the premises applied at the time of intervention approval and during reporting cycles. Aggregate analyses of (validated) self-evaluation reports enable cross-sectoral and cross-regional comparisons of performance as well as reporting at the corporate levels and to the Boards.

**Demand-driven self-evaluation.** A variety of evaluation activities are undertaken in response to specific donor, client or internal demands, or as an element of operational or research work—for example, retrospective studies of various products and instruments, trust fund evaluations and impact evaluations to assess the impact of activities and interventions. Demand-driven self-evaluations are structurally embedded in managerial processes. However, they are often either conducted, managed or commissioned from external consultants by functionally independent units within the institution.

On mandatory self-evaluation matters, the World Bank's Operations Policy and Country Services (OPCS) department, led by the OPCS Vice President, is the primary group (and counterpart unit to IEG) responsible for the World Bank's self-evaluation system. OPCS represents World

Bank Management in official interactions with IEG. While IEG engages with other departments on specific studies (e.g. the Human Development department for evaluations involving the education sector), it is OPCS that is responsible for M&E arrangements for World Bank operations. IEG interacts on evaluation matters with the IFC and MIGA, but a discussion of those engagements and products is outside of the scope of this paper.

Self-evaluation—the formal, written assessment of a project, programme or policy by an entity engaged in that activity—lies at the heart of the World Bank Group’s results measurement system and has been used to assess the outcomes of investments for 40 years. In 2016, a major evaluation from IEG ‘A Report on the Self-Evaluation Systems of the World Bank Group’ assessed how well the World Bank Group’s self-evaluation systems served their expected purposes.

That report summarised the main elements of the self-evaluation system as follows:

The Bank began requiring all operating departments to prepare self-evaluation project completion reports in 1976. Those early reports were subject to review by the evaluation department (now known as IEG) before being submitted to the Board. ...A decline in the development effectiveness of Bank projects in the early 1990s spurred a number of changes. The ICR was introduced with validation by IEG after submission to the Board rather than before, resulting in rating differences between the ICR and IEG’s validation. The Quality Assurance Group was established (and later disbanded) to evaluate the quality at entry, quality of supervision, and overall portfolio performance. Over time, the independent evaluation function has worked closely with Operations Policy and Country Services (OPCS) to harmonise rating systems, adjust ratings criteria, and introduce new self-evaluation products. Currently, World Bank carries out self-evaluations which IEG validates for all IBRD/IDA operations regardless of funding size and all recipient executed trust funds above \$5 million (with a few exceptions). The evaluations assess the project against the original project objectives and any subsequent formal revisions and rate outcomes based on criteria for relevance, effectiveness, and efficiency. Risk to development outcome, Bank performance, and borrower performance are also assessed and rated. In addition, IEG separately assesses and rates project M&E and the quality of the self-evaluation. IEG also writes Project Performance Assessment Reports on a purposefully selected share of projects, currently around 15 percent. (Independent Evaluation Group, 2016, pp. 4–5).

One area where the World Bank and IEG have had influence in promoting M&E tools involves the requirement of World Bank projects to use theories of change and results frameworks in project design, monitoring and evaluation activities. Beginning in 1997, the World Bank adopted the logical framework, or logframe, as the core reference document used throughout the entire project management cycle. At the time, all Project Appraisal Documents for investment operations were required to create a logframe to plan the project goals, objectives and outputs, among other things. The logframe approach was based on what was created in 1969 for the US Agency for International Development and later in the bilateral community (Team Technologies, 2005).

In the 2000s, the World Bank moved from the logframe to the results framework, to show a greater articulation of the outcome indicators, means of verification and so on. Finally, in 2017, the World Bank required that, in addition to the results frameworks, all investment projects would require an accompanying theory of change to visually explain the expected causal pathway. With all of these tools required in World Bank projects, while no known study has been made, the World Bank has very likely had a leading role in the mainstreaming of the use of these M&E tools in development projects and programmes globally.

Because World Bank-financed projects require results frameworks and theories of change for project designs and monitoring purposes, a result has been more familiarity and systematic use within counterpart government offices and project implementing agencies.

IEG does not conduct impact evaluations, but impact evaluation is a common and important type of demand-driven evaluation conducted in the World Bank. Much impact evaluation is financed through trust funds provided by a range of donors. The methodological approaches taken through these impact evaluations typically involve randomised control trials or quantitative quasi-experimental techniques. Though not the only teams working on impact evaluation, the largest are the Development Impact Evaluation Initiative (DIME), the Africa Gender Lab, the Strategic Impact Evaluation Fund (SIEF) and the Health Results Innovation Trust Fund. The impact evaluations operating through these teams are not undertaken for systematically corporate accountability purposes (though they do fulfil some accountability purposes). Instead, they are more typically used for learning purposes for the projects and sectors under study. In this way, the impact evaluations conducted by these teams have more of a research orientation. IEG does not validate impact

evaluations, nor does IEG join in impact evaluations with these teams (though IEG has carried out a small set of systematic reviews using impact evaluation results).

When quantitative impact evaluations became prevalent in the early 2000s, following on the lessons coming out of important impact evaluations around programmes such as Mexico's PROGRESA, IEG joined with others to try to understand the changing evaluation environment. Beginning in 2006, IEG held the secretariat for NONIE, the Network of Networks for Impact Evaluation, since disbanded. NONIE was formed by the OECD and DAC Evaluation Network, UNEG, ECG and the International Organisation for Cooperation in Evaluation (IOCE)—a network drawn from the regional evaluation associations. At a time when quantitative impact evaluations emphasising quantitative counterfactual analysis were gaining steam, NONIE worked to develop a common understanding of the meaning of impact evaluation and approaches to conducting impact evaluation. The working paper, 'Impact Evaluations and Development: NONIE Guidance on Impact Evaluation', was developed through the network's efforts (Leeuw & Vaessen, 2009).

## IEG IN THE AMERICAS AND GLOBALLY: INFLUENCES, ORGANISATIONAL STRUCTURES, ACTIVITIES

As noted in the introduction, IEG operates globally, making the task of focusing on the Americas somewhat difficult. But this section is meant to locate IEG within the Americas and provide examples of how IEG has been influenced by those in the region, and how IEG has influenced those in the region.

### *Influences on IEG in the Americas: GAO*

IEG has not only served as an influencer but has also been influenced by other groups. Most notably, in the Americas, the United States General Accounting Organisation, renamed the Government Accountability Office (GAO) in 2004, the main audit body of the US Congress, was the agency most involved with early 1970s formation of OED

as an independent evaluation group (which became IEG).<sup>2</sup> The GAO had experience in assessing US ‘Great Society’ and other programmes over the years and had lessons to share on how evaluations could be conducted. Evaluation and research activities on the World Bank’s projects had certainly occurred over the years prior to OED forming. Over a period of 1968–1973, experiences were gained and thinking evolved and what could be. Working alongside the World Bank’s Board members, World Bank President McNamara, and Bank Management and staff involved with evaluation functions, the GAO had a role in shaping OED’s initial organisational structure, methodological approaches and types of evaluations (Willoughby, 2002).

In the 1980s to mid-1990s, the GAO’s Programme Evaluation and Methodology (PEMD) was directed by Eleanor Chelimsky (1980–1994). She and PEMD were influential in evaluation work and methodological approaches for the US government, and also influential in the broader professional evaluation community and academia. GAO’s evaluation and methodology lessons were of interest to IEG leadership and staff, and informal exchanges occurred with Chelimsky and members of PEMD staff. Starting in the late 1990s as PEMD transitioned within GAO and Chelimsky retired, three senior PEMD staff members joined IEG in senior and management roles, bringing their GAO expertise with them and embedding it within IEG’s work and environment (Author interviews with Morra and Rist, 2009).

### *IEG Influences: Evaluation Functions in Global and Regional Multilateral Development Banks and Networks*

The World Bank was the first multilateral development bank, and IEG was the first independent evaluation body in a multilateral development bank. Most of the multilateral bodies have modelled both their organisational structures and functions, including the self-evaluation arrangements, on the World Bank’s approach. Among others, global agencies that have adopted (and adapted) IEG’s model, having set up independent evaluation offices within them, include the Asian Development Bank (ADB), African Development Bank (AfDB) and the International Fund for Agricultural Development (IFAD). In the Americas, the Office of Evaluation

<sup>2</sup> In 2004, the GAO was renamed the Government *Accountability* Office to better encompass its work beyond auditing.



and Oversight (OVE) of the Inter-American Development Bank (IDB) was established in 1999 (after separating its earlier functions and ties to IDB Management) and has shaped elements of its arrangements on IEG's experiences, with adjustments to IDB's own needs. Another example is the smaller Banco Centroamericano de Integración Económica, BCIE (Central American Bank for Economic Integration, CABEI) with its Oficina de Evaluación (ODE), or Office of Evaluation. ODE was created in 2009 and placed under the Office of the Chief Economist and became an independent evaluation office in 2012 and began its direct reporting relationship with BCIE's Board of Directors. Beyond the organisational structures, the independent evaluation offices of regional multilateral development banks have been headed at times over the last decade by former IEG directors (once at the IDB and twice at the ADB).

*IEG Influences: Role in International Agency Networks  
and Initiatives*

IEG and IEG staff are engaged to varying degrees in international agency networks and initiatives involving aid effectiveness and supporting better evaluation. In and through these and other groups, IEG has been involved in a range of efforts such as the High-Level Fora on Aid Effectiveness in Rome (2003), Paris (2005), Accra (2008) and Busan (2011); Millennium Development Goals (MDGs); Sustainable Development Goals (SDGs); and others.

Two key networks with which IEG works regularly are ECG and DAC EvalNet.

*ECG*

IEG had an influential role in the 1996 establishment of the ECG, a network for multilateral organisations with evaluation functions. In 1996, a Development Committee Task Force (officially known as the Joint Ministerial Committee of the Boards of Governors of the World Bank and International Monetary Fund on the Transfer of Real Resources to Developing Countries) issued a report that called for, among other things, the harmonisation of evaluation methodologies, performance indicators and criteria. Heads of the evaluation offices of the following MDBs came together as founding members of ECG: AfDB, ADB, European Bank for Reconstruction and Development (EBRD), IDB and the IEG at the World Bank Group. The ECG's mandate focused on harmonisation of

evaluation principles, standards and practices and facilitating the involvement of borrowing member countries in evaluation. ECG also addresses issues related to accountability, learning from past experiences, sharing these lessons and strengthening their use. Since its establishment, other international financial institutions have joined the ECG or been involved through observer status, including the UNEG (World Bank, 1996).

The ECG members have established a secretariat, have a rotating chairmanship arrangement and meet twice per year. In 2012, the ECG developed the ‘ECG Big Book on Good Practice Standards’ to support ECG’s mandate on harmonisation and on improving understanding of evaluation practices. In addition to harmonisation, ECG members also work together on matters of professionalism, lessons learning and cooperation.

### *DAC EvalNet*

IEG is one of the nine multilateral agencies joining 37 bilateral members of DAC EvalNet, which is the OECD DAC on Development Evaluation (EvalNet). DAC EvalNet members work together on important evaluation topics and the contributions of this group to the broader evaluation community are vast, with many resources available in multiple languages. Among them are the DAC Criteria for Evaluating Development Assistance, Glossary of Key Terms in Evaluation and Results-Based Management, and Evaluation Systems in Development Co-operation (covering recent trends and challenges in the evaluation systems of members). In 2019, this group approved the Revised Evaluation Criteria, updating an important guide developed originally in 1991 and defined more in 2002 (OECD, DAC, 2019).

### *IEG Influences: Transparency, Open Access and Outreach*

IEG approaches issues of transparency, open access and outreach in different ways, with some examples provided below.

In 2010, the World Bank took an important step for government partners, citizens, researchers and others in establishing the World Bank Policy on Access to Information, building on earlier disclosure policies. The policy is based on five principles: maximising access to information; setting out a clear list of exceptions; safeguarding the deliberative process; providing clear procedures for making information available; and recognising requesters’ right to an appeals process (World Bank, 2010). IEG

already had a mix of disclosure policies, but in 2011 articulated a single Access to Information Policy to promote greater transparency and consistency of access across IEG units. IEG's evaluations are available on its website, including validation reports for individual World Bank projects.

IEG has worked with World Bank Management counterparts to have a means to track actions coming out of recommendations in evaluations. This occurs through the management action record (MAR) system. At the conclusion of an evaluation, Management develops, in consultation with IEG, an action plan for implementation against recommendations in IEG's evaluations. These actions are tracked and assessed over a four-year period and are publicly disclosed in the MAR system, which is available on IEG's website. At the time of the publication of this book, the MAR system is undergoing a reform to improve its utility.

IEG devotes a good deal of resources to dissemination and outreach. Recently, IEG has moved towards developing shorter evaluation reports, videocasting evaluation launch events, blogging about findings from evaluations, and using social media to professional evaluator and development communities. In 2017, through blog posts and collaborations with the head of DAC EvalNet and others in the professional development community, IEG's director general promoted a re-visiting of the DAC Evaluation Criteria. This prompted an open international review leading to the updated criteria (OECD, DAC, [2019](#)).

### ***IEG Influences: Evaluation Capacity Development (ECD)***

IEG has been involved with ECD activities over the years, with ebbs and flows in the level of activity. Early work on ECD involved documenting its own approaches and methodologies and sharing them with others and learning from others to do better in the nascent field of evaluation. IEG began doing limited technical advisory work in the 1990s with government agencies wanting to set up arrangements and systems for both monitoring and evaluation functions. IEG became involved with additional advisory work, often with other World Bank operational teams and other agencies such as the IDB, with counterpart countries. In the late 1990s, and especially during the first decade of the 2000s, IEG staff authored and commissioned research to document the experiences of countries in setting up their M&E systems. Additional research activities involved developing publications on evaluation and results-based management methodologies. Support in developing and strengthening

professional associations, or voluntary organisations of professional evaluators (VOPEs), was another area of focus in the 2000s. And finally, IEG developed and co-sponsored important training and other capacity development programmes.

### *M&E System Advisory*

In 1994, the World Bank's strategy for evaluation capacity development (ECD) was established with the goal of helping borrowers strengthen their M&E systems as an integral part of sound governance. This was a consequence of the so-called Wapenhans Report, or the report on *Effective implementation: key to development impact*, named after the World Bank vice president who headed the study, Willi Wapenhans, which focused on how to have improved performance and quality of World Bank operations (World Bank, 1992). A task force was set up to develop this and it recommended 'mainstreaming' the Bank's support for evaluation capacity. The committee that developed the strategy was led by the World Bank's central operations team and involved issues around policy, operations, budget and so on. While the vision was commendable and areas of success, the implementation was judged to be uneven (World Bank, 2002).

World Bank Management renewed its effort to mainstream ECD in fiscal year 2004. M&E system advisory work over next ten years had a period of strong growth, particularly in the Latin America and Caribbean region and Africa region. Following a 2013 reorganisation of the World Bank and a range of other organisational, staffing and other changes, the momentum diminished among World Bank operational teams and became more fragmented. IEG, too, reoriented to other ECD efforts, including working on the CLEAR Initiative, which will be described below.

Because of the many actors involved and usually lengthy periods of time involved in developing and nurturing M&E systems, it is difficult to delineate the role of IEG from other actors. Normally, IEG has played just one part in M&E system development. Over the years, IEG has acted in concert with World Bank operational counterparts (often those working on public sector administration matters, but sometimes in sectoral areas such as health) in advisory functions. In the last 20 years, especially in the first decade, IEG provided advisory services with the Latin America Region staff of the World Bank. The region is highly regarded in the evaluation community, especially in the middle-income countries, of having

the more successful experiences with establishing laws, policies, guidelines, processes, technical staff, leaders and so on. Often cited examples include Colombia, Chile and Mexico, among others.

Among an early documented example of government interest in evaluation system development, was in Colombia, and came as a result of the government seeing the power of information coming from an IEG evaluation. At the request of the Minister of Finance, IEG was asked to advise the newly appointed government lead on evaluations. Later, IEG's director general, Yves Rovani, and his deputy, Pablo Guerrero, along with Eleanor Chelinsky of the GAO, participated in a high-level conference in Colombia, reviewing and advising (Grasso et al., 2003).

A summary of what occurred was captured by Pablo Guerrero (Guerrero Ortiz, 1999, p. 2).

In Colombia, in 1991, the government of President Cesar Gaviria commissioned a high-level study to assess how Colombia should institutionalise an evaluation function. The government had been exposed to an external evaluation of investments in the key power sector, showing that Colombia could have improved substantially the outcome and services of multi-billion dollar investments in the sector. The high-level study was commissioned to study the feasibility of introducing evaluation to cover major public sector expenditures. While doing the study, the government quickly recognised that to institutionalise the function would require political consensus bridging the interests of all concerned parties. To do this, the government commissioned five independent evaluations of large, controversial public sector programmes. These programmes covered a period which had contained various political regimes, and included a hydroelectric, an urban rapid transit, a papermill, a community welfare, and an international communications project (DNP 1992). The aim of the evaluations was to enrich dialogue among the various interest groups by providing them with specific evaluation results comparing what had been achieved with what could have been done, had those evaluations been available earlier in the decision-making process. The results of these evaluations were discussed in a high-level conference where past and present government decision-makers, international experts, and political interest groups participated. The conclusion of the conference led to a recommendation that the Colombian constitution, which was being redrafted at the time, should include an article making public sector evaluation mandatory. Thus, through leadership, stakeholder involvement, and opportunism, a legal foundation for evaluation was established in Colombia.

In the 2002 'IEG Annual Report on Evaluation Capacity Development', the author documented ECD activities receiving World Bank support (from operational teams and, in some cases, IEG) in 21 countries, including eight in the Africa Region, five in Latin America and the Caribbean, four in Europe and Central Asia, two in South Asia and one each in East Asia & the Pacific, and the Middle East & North Africa regions. The countries were Albania, Argentina, Bangladesh, Bolivia, Brazil, Chile, Egypt, Ethiopia, Ghana, Honduras, India (Andhra Pradesh), Kyrgyz Republic, Madagascar, Malawi, Mozambique, Niger, the Philippines, Poland, Romania, Tanzania and Uganda.

For the five countries in Latin America—Argentina, Bolivia, Brazil, Chile and Honduras—the authors noted that Argentina, Chile and Colombia had engaged in ECD efforts for some time, with some assistance from the Bank and other donors. The report further mentioned that the World Bank's Latin America and the Caribbean Region was preparing a regional M&E strategy and would likely to lead to greater regional efforts on ECD.

Jumping ahead to 2010, the Latin American and Caribbean Region (with the IDB and Government of Spain) documented important M&E systems work in the region that was shared in the region's fifth M&E Conference on this topic. The conference topic was 'Challenges in Monitoring and Evaluation: An Opportunity to Institutionalise M&E Systems'. The event was hosted by the Colombian National Planning Department (Departamento Nacional de Planeación), with the support of the government of Spain. IEG provided a supportive role in the conference and report.

The conference is an example of the type of convening role that the World Bank (including IEG) plays with other partners in bringing together key stakeholders to share, learn and network. The report is important because it documented the experiences in Latin America, but also of other countries including Canada, Sri Lanka, Spain and South Africa. It further covered topics such as challenges to evidence-based decision-making; institutional arrangements for M&E Systems; capacity building for monitoring and evaluation and academia-government-civil

society partnerships; and institutional arrangements and related policies to ensure the quality, access and use of performance.

*Research: M&E Systems and Methodological Topics*

One of IEG's (and the World Bank more generally) most enduring sources of influence involves its research activities, in particular with its open access policy. In the late 1990s into the current decade, IEG has produced research materials in two broad topics: M&E systems and evaluation methods. Many of these publications have been translated into a range of languages.

The *M&E systems series* and papers documented experiences and lessons from countries across the globe, ranging from Colombia, to South Africa, to Australia. Authors with direct expertise on the systems in those countries were commissioned to prepare the reviews. The studies typically provide information about the evolution of system development and features of the systems (such as organisational structures, laws, policies and so on). They offer a comparative look across countries and contributed greatly to the literature base on this topic. In the Americas, studies have been done for Canada, Chile, Colombia (including subnational), Mexico, the United States and comparative studies of Latin American countries (Argentina, Chile, Colombia, Costa Rica and Uruguay).

Other studies in this M&E systems series provided lessons that could be applied in all contexts, such as Keith MacKay's 'How to Build M&E Systems to Support Better Government' or Marc Robinson's 'Connecting Evaluation and Budgeting'.

IEG authors, sometimes with other World Bank staff, produced texts on *methodological topics* that proved highly popular. 'Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners', from IEG's Ray Rist and the World Bank's Jody Zall Kusick, has been a top 25 download on the World Bank's website of all publications (including many years of the World Bank's flagship publication, the *World Development Report*). 'The Road to Results: Designing and Conducting Effective Development Evaluations', from Rist and IEG's Linda Morra, is a top 50 download. These books have been translated from Spanish to Vietnamese to Korean to other languages, often on request and by a requesting counterpart agency in a country, such as South Korea. The metadata from the World Bank's Open Knowledge Repository and the requests to translate the books—often used in

global university classrooms—is evidence of the demand for them and value of the World Bank investing in these types of publications. Beyond this, over the years IEG produced a series of popular briefs, also available in a number of languages, such as ‘Monitoring and Evaluation: Some Tools, Methods, and Approaches’ and ‘Writing Terms of Reference for an Evaluation: A How To Guide’.

Starting in 2010, IEG has initiated (with partners) two main capacity development programmes: the International Programme for Development Evaluation Training (IPDET) and the CLEAR Initiative.

The IPDET was founded in 2001 by the World Bank, with more donors supporting the programme over the years. IPDET is a residential executive training programme built around evaluation topics, aimed at evaluation practitioners, programme managers, policymakers and others. From 2001 to 2016, IPDET was offered at Carleton University in Ottawa, Canada, and trained more than 3.500 participants from 125 countries.

In 2017, IPDET 2.0 moved to Bern, Switzerland, and has had programmes in 2018 and 2019 in a partnership among the Center for Continuing Education at the University of Bern (ZUW), the Center for Evaluation at the Saarland University in Germany (CEval) and IEG. ZUW and CEval are two key European evaluation hubs, both offering the longest-standing demand-oriented master’s and continuing education programmes in evaluation. In 2019, IPDET’s core course faculty included two trainers—one from South Africa and one from Mexico—from the CLEAR Initiative.

The IPDET programme currently offers a one-week core course on fundamental evaluation topics, followed by two weeks of specialised shorter workshops offered in parallel. The courses are taught by renowned international faculty drawn from organisations around the world.

At the time IPDET was begun, there were no-such regularly offered residential programmes with a focus on development evaluation. And the overall supply of evaluation training programmes was, generally, unpredictable with limited offerings. Almost 20 years after IPDET started, the supply of evaluation training programmes across the globe has multiplied, with many programmes taking inspiration from IPDET.

In the Americas, in addition to IPDET operating in Ottawa for 15 years, an off-shoot of IPDET services French-speaking Canada (and also parts of the Caribbean and Africa). PIFED (Le Programme International de Formation en Evaluation du Développement) is the French



language version of the IPDET programme, but operates independently from IPDET in Bern. ‘École Nationale d’Administration Publique’ (ENAP) in Québec City has run the programme since 2011, with funding from the Government of Canada.

In commencing the CLEAR Initiative (the Regional Centers for Learning on Evaluation and Results), donors were motivated to support a new programme that would promote and support *locally driven and owned* ECD efforts. The CLEAR Initiative began in 2010 as a programme initiated by IEG and with multilateral development banks and bilateral agencies including the following founding members: AFDB; ADB; Australian Government—Department of Foreign Affairs and Trade; Belgian Federal Public Service for Foreign Affairs-International Trade and Development Cooperation; IDB; the Rockefeller Foundation; Swedish International Development Cooperation Agency (Sida); Swiss Agency for Development and Cooperation (SDC); United Kingdom Department for International Development (DFID); and the World Bank.

CLEAR has six regional ECD centres (and affiliates) competitively selected and established within universities across the globe. CLEAR’s six centres are: Anglophone Africa (in South Africa, with an affiliate in Ghana), Brazil and Lusophone Africa (in Brazil), East Asia (in China), Francophone Africa (in Senegal), Latin America and the Caribbean (in Mexico), and South Asia (in India, with an affiliate in Pakistan). The centres work within their respective regions offering these main products and services: training, technical assistance, knowledge products and knowledge sharing. IEG serves as the CLEAR Global Hub and manages donor funding, offers strategic guidance and promotes learning and collaboration within CLEAR and the global M&E community.

CLEAR centres receive grants through the programme and also depend on resources from their own universities and other sources of funds. In the fiscal year ending June 30, 2018, CLEAR delivered capacity-building trainings to 28,277 individuals from 66 countries. But CLEAR centres do much more than training. They have advised government agencies on setting up and improving M&E systems, developed research projects, joined and led VOPE initiatives, among many other services (The CLEAR Initiative, 2018).

In the Americas, the two centres located in Mexico and Brazil also work closely with the IDB’s OVE (a donor to CLEAR) in programme planning. The CLEAR Center in Mexico, located at Centro de Investigación y Docencia Económicas (CIDE), began operations in 2012

and has over 20 full-time staff plus a network of over 50 professors, consultants and experts from Latin America and abroad. This centre has delivered over 60 custom and open-enrolment courses in 14 Latin American countries. They have conducted more than 20 evaluations and published over a dozen books and other knowledge products. An example of one of the centre's feature books is, 'Panorama de los sistemas nacionales de monitoreo y evaluación en América Latina' (*Overview of the national M&E systems in Latin America*). In 2017, the centre was selected by IFAD to receive a grant to develop and run an international training programme in rural M&E. The programme, PRiME (Programme in Rural M&E), offers training in this topic in Spanish, French and English and has trained nearly 200 people from 75 countries, with repeat programmes planned. The PRiME programme continues and IFAD provided a second grant for expansion and sustaining the programme in 2020.

Corresponding with the growth of VOPEs globally, this centre organised the first Evaluation Week in Mexico. In 2017, the centre (and other partners in Latin America) expanded this event to reach other countries in the Latin America. In 2019, in conjunction with the inaugural CLEAR gLOCAL Evaluation Week, more than 100 institutions hosted events and we reached an audience of over 10,000 people in 14 Latin American countries.

The CLEAR Center for Brazil and Lusophone Africa is hosted by the Getulio Vargas Foundation (FGV) at the São Paulo School of Economics and began operations in 2015. This centre has worked closely with national and subnational governments in Brazil to improve programme results and inform budgetary decisions. Starting in 2018, the centre began working with the Brazilian state of Espírito Santo to design its own M&E system and strengthen their organisational capacities. The centre works not only with governments, but also civil society organisations, private sectors and foundations. For example, the team conducted an evaluative study on labour market and skill gaps in São Paulo that was sponsored by JPMorgan's 'New Skills at Work'. The centre works closely with the Brazilian M&E Network (RBMA) to co-organise conferences, support fundraising and provide in-kind resources for outreach activities and trainings. The centre also prioritises offering its services in states in Northeast Brazil, the economically poorest region of Brazil.

More recently, the centre has begun to collaborate with the CLEAR Anglophone and Francophone centres in Africa to have a cohesive

working arrangement in Africa and to better meet the needs of Portuguese-speaking counterparts on the African continent.

### *Evaluation Networks and VOPEs*

As mentioned earlier, over the years, IEG has participated in various networks, including ECG, DACEvalNet and NONIE. IEG is also a formal observer to the UNEG. While arrangements are more informal, IEG has collaborated on various initiatives with International Initiative for Impact Evaluation (3ie), EvalPartners and regional VOPEs such as the African Evaluation Association (AfrEA), American Evaluation Association (AEA), European Evaluation Society (EES) and many country-level associations.

A notable network started in 2005 by the World Bank, the Prodev programme of the IDB and a growing group of Latin American countries, was REDLACME (the Regional Monitoring and Evaluation Network for Latin America). Its objectives involved bringing together mostly government counterparts and donors to promote dialogue, share experiences, strengthen institutional capacity and stimulate collaboration related to the institutionalisation of M&E Systems at the national and subnational levels. The World Bank and the IDB played a catalytic role in developing an online platform for members to share information about best practices, events, evaluation methods and other resources.

Working with other partners and meant to be a counterpart initiative to IPDET, IEG initiated the effort on the International Development Evaluation Association (IDEAS). IDEAS began in September 2002 to meet the need for a global professional association for internationally active development evaluators. This important network was started at a time before EvalPartners and others promoting VOPEs really got underway. In almost 20 years, the global evaluation community has changed greatly with many new and more robust national and regional associations. Though supportive of IDEAS, IEG's formal role with IDEAS is more limited than in its origin. IDEAS today is a global professional evaluation association which focuses on international sustainable development. IDEAS's main efforts involve promoting the profession, fostering capacity development, and improving and advancing theories, methods and use of evidence.

Finally, at the time of the completion of this chapter, IEG has launched—with a range of partners—a major new ECD effort, the Global Evaluation Initiative (GEI). GEI is a Global Partnership Programme with a multidonor trust fund that is managed by IEG. Initial contributing

donors—more specifically, the evaluation offices within the respective agencies—include Sida, the Ministry of Foreign Affairs of the Netherlands, the German Federal Ministry for Economic Cooperation and Development (BMZ), the Swiss Agency for Development and Cooperation (SDC), the Ministry of Foreign Affairs of Finland, the Ministry of Foreign Affairs of Denmark, Global Affairs Canada, UNDP, IFAD and IDB.

GEI will connect donors, governments, citizens and experts to support countries in strengthening monitoring and evaluation frameworks and capacities. At the donor level, GEI will address issues of fragmentation, to better connect, coordinate and finance ECD programmes and areas of support. The GEI programme will work globally and in multiple languages and serve to connect and support technical partners and others working in ECD—such as the CLEAR Initiative, IPDET, BetterEvaluation, 3ie, VOPEs and many others. GEI's efforts are centred around four main activity components including technical advisory services (including M&E system diagnostics and situational analyses), training, knowledge generation and knowledge sharing. Ultimately, GEI will work to promote M&E frameworks, processes, systems among individuals and organisations served by GEI globally, and contribute to improving the 'enabling environment' for better M&E in countries where GEI partners will be present. While GEI is a global programme, GEI will have a strong presence in the Americas through both its donor partners (especially IDB and BMZ) and technical partners—starting with the CLEAR centres in both Brazil and Spanish-speaking Latin America—bringing global lessons to the Americas and sharing lessons from the Americas throughout the globe.

## CONCLUSION

The paper has summarised how IEG was set up and functions within the World Bank Group, and through that elaboration, provided examples of IEG's contributions in the Americas and globally in these important areas.

IEG has served as a *model for other agencies seeking to set up independent evaluation departments*, particularly other global and regional international financial institutions (IFI). IEG's organisational structure, mandate and functions have been replicated and customised in organisations such as in independent evaluation offices in the IDB, ADB, AfDB, IFAD and others.

The *self-evaluation system* in the World Bank Group, of which IEG has a role, has been served as a basis for how IFIs and other groups conduct their evaluation efforts. Because World Bank-financed projects require results frameworks and theories of change for project designs and monitoring purposes, a result has been more familiarity and systematic use within counterpart government offices and project implementing agencies.

*Convening and partnering in a number of collaborative* efforts to promote the exchange of ideas and furthering of shared goals has helped to shape the field of evaluation. Notable examples of areas where IEG has contributed include the following: Evaluation Cooperation Group (ECG), which began in 1996 and continues; Network of Networks for Impact Evaluation (NONIE) (2006–2009); and the International Development Evaluation Association (IDEAS), which began in 2002 and continues, and is a professional association focusing on evaluation in developing countries in partnership with UNDP, OECD’s DAC Network on Development Evaluation (DAC-EvalNET), and the UNEG. In the Americas, REDLACME was organised with IDB and countries in Latin America to promote dialogue and sharing in that region.

*Research on evaluation topics* has been an important area of IEG’s work over the years, including methodologies, standards, curriculum, M&E systems and professionalisation.

*Technical advisory to countries on developing M&E systems* has been an area where IEG has applied many lessons from its research on M&E systems and used information gained through advisory work to feed into new research.

*Evaluation capacity development (ECD) efforts*—notably the programme that started almost 20 years ago, IPDET, and in the last ten years, the CLEAR Initiative—have had global influence in the growing field of ECD. Through the efforts of these two programmes, many individuals and organisations have been influenced by these two programmes. Finally, through the establishment of GEI in 2020, it is expected that IEG—along with many donors and other partners—will contribute to less fragmentation and better results in promoting and strengthening ECD in countries across the globe at individual, organisational and enabling environment levels.

## LIST OF ABBREVIATIONS

3ie	International Initiative for Impact Evaluation
ADB	Asian Development Bank
AEA	American Evaluation Association
AFDB	African Development Bank
AfrEA	African Evaluation Association
BCIE	Banco Centroamericano de Integración Económica
BMZ	German Federal Ministry for Economic Cooperation and Development
CEval	Center for Evaluation (at the Saarland University in Germany)
CIDE	Centro de Investigación y Docencia Económicas
CLEAR	Centers for Learning on Evaluation and Results
CODE	Committee on Development Effectiveness
DAC	Development Assistance Committee
DGE	Director General of Evaluation
DFID	Department for International Development (of the United Kingdom)
DIME	Development Impact Evaluation Initiative
EBRD	European Bank for Reconstruction and Development
ECD	Evaluation Capacity Development
ECG	Evaluation Cooperation Group
EES	European Evaluation Society
ENAP	École Nationale d'Administration Publique' (ENAP)
FGV	Getulio Vargas Foundation
GAO	Government Accountability Office
GEI	Global Evaluation Initiative
IDB	Inter-American Development Bank
IBRD	International Bank for Reconstruction and Development
ICR	Implementation Completion and Results Report
ICRR	Implementation Completion and Results Report Review
ICSID	International Centre for Settlement of Investment Disputes
IDA	International Development Association
IDEAS	International Development Evaluation Association
IEG	Independent Evaluation Group

IFAD	International Fund for Agricultural Development
IFC	International Finance Corporation
IFI	International Financial Institution
IOCE	International Organisation for Cooperation in Evaluation
IPDET	International Programme for Development Evaluation Training
MAR	Management Action Record
MIGA	Multilateral Investment Guarantee Agency
NONIE	Network of Networks for Impact Evaluation
OECD	Organisation for Economic Co-operation and Development
OED	Operations Evaluation Department
OVE	Office of Evaluation and Oversight
OPCS	Operations Policy and Country Services
PEMD	Programme Evaluation and Methodology
PIFED	Le Programme International de Formation en Evaluation du Développement
PPBS	Planning, Programming, and Budgeting System
PRiME	Programme in Rural M&E
REDLACME	Regional Monitoring and Evaluation Network for Latin America
SDC	Swiss Agency for Development and Cooperation
Sida	Swedish International Development Cooperation Agency
SIEF	Strategic Impact Evaluation Fund
UNEG	United Nations Evaluation Group
VOPE	Voluntary Organisations of Professional Evaluators
WBG	World Bank Group
ZUW	Center for Continuing Education (at the University of Bern)

## REFERENCES

- Grasso, P. G., Wasty, S., & Weaving, R. (2003). *World Bank operations evaluation department: The first 30 years*. World Bank. <https://openknowledge.worldbank.org/handle/10986/15128>. Accessed on 14 December 2019.

- Guerrero Ortiz, R. P. (1999). Comparative insights from Colombia, China, and Indonesia. World Bank, Operations Evaluation Department. *Evaluation Capacity Development*, 5. World Bank. <http://documents.worldbank.org/curated/en/640511468018571381/Comparative-insights-from-Colombia-China-and-Indonesia>. Accessed on 14 December 2019.
- Independent Evaluation Group. (2016). *Behind the mirror: a report on the self-evaluation systems of the World Bank Group*. Corporate and process evaluation. World Bank. <http://documents.worldbank.org/curated/en/902331469736885125/Behind-the-mirror-a-report-on-the-self-evaluation-systems-of-the-World-Bank-Group>. Accessed on 14 December 2019.
- Leeuw, F. L., & Vaessen, J. (2009). *Impact evaluations and development: NONIE guidance on impact evaluation*. Network of networks on impact evaluation. World Bank. <http://documents.worldbank.org/curated/en/411821468313779505/Impact-evaluations-and-development-NONIE-guidance-on-impact-evaluation>. Accessed on 14 December 2019.
- Lopez-Acevedo, G. C., Rivera, K. L., Lima, L., & Hwang, H. (2010, June). *Challenges in monitoring and evaluation: An opportunity to institutionalize M&E systems* (Working paper of the World Bank and Inter-American Development Bank No. 55853).
- Morra Imas, L. G., & Rist, R. (2009). *The road to results: Designing and conducting effective development evaluations*. The World Bank. <https://hubs.worldbank.org/docs/imagebank/pages/docprofile.aspx?nodeid=11665295>. Accessed on 14 December 2019.
- OECD, DAC. (2019, December). *Evaluation criteria: Adapted definitions and principles for use*. Report DCD/DAC, (2019)58, DAC Meeting. [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DCD/DAC\(2019\)58/FINAL&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DCD/DAC(2019)58/FINAL&docLanguage=En). Accessed on 14 December 2019.
- Team Technologies. (2005, January). *The logframe handbook: A logical framework approach to project cycle management* (Working paper of the World Bank). <http://documents.worldbank.org/curated/en/783001468134383368/The-logframe-handbook-a-logical-framework-approach-to-project-cycle-management>. Accessed on 14 December 2019.
- The CLEAR Initiative. (2018). *CLEAR annual report 2018*. [https://www.theclearinitiative.org/sites/default/files/2019-01/CLEAR\\_Annual\\_Report\\_2017-18.pdf](https://www.theclearinitiative.org/sites/default/files/2019-01/CLEAR_Annual_Report_2017-18.pdf). Accessed on 14 December 2019.
- Willoughby, C. (2002, September). *The first experiments in operations evaluation: roots, hopes, and gaps* (Working paper of Operations Evaluation Department [OED] No. 24915). <http://documents.worldbank.org/curated/en/962861468158708263/The-first-experiments-in-operations-evaluation-roots-hopes-and-gaps>. Accessed on 14 December 2019.



- World Bank. (1992). *Effective implementation: Key to development impact (Wapenbans Report)*. World Bank. <https://hubs.worldbank.org/docs/imagebank/pages/docprofile.aspx?nodeid=736775>. Accessed on 14 December 2019.
- World Bank. (1996). *Serving a changing world—Report of the task force on multilateral development banks*. World Bank. <https://hubs.worldbank.org/docs/imagebank/pages/docprofile.aspx?nodeid=5436236>. Accessed on 14 December 2019.
- World Bank. (2002). *Annual report on evaluation capacity development—2002*. World Bank. <http://documents.worldbank.org/curated/en/609421468764051323/Annual-report-on-evaluation-capacity-development-2002>. Accessed on 14 December 2019.
- World Bank. (2003). *Independence of OED*. OED Reach. World Bank.
- World Bank. (2010). *The World Bank policy on access to information*. World Bank.
- World Bank. (2018). *Independent evaluation group mandate*. World Bank.
- World Bank. (2019a). *World Bank annual report 2019*. World Bank. <https://openknowledge.worldbank.org/bitstream/handle/10986/32333/9781464814709.pdf>. Accessed on 14 December 2019.
- World Bank. (2019b). *Independent evaluation group annual report FY19: transforming evidence into better outcomes*. World Bank. <https://hubs.worldbank.org/docs/imagebank/pages/docprofile.aspx?nodeid=31446622>. Accessed on 14 December 2019.
- World Bank. (2019c). *World Bank group evaluation principles*. World Bank. <https://ieg.worldbankgroup.org/sites/default/files/Data/reports/WorldBankEvaluationPrinciples.pdf>. Accessed on 14 December 2019.

PART IV

---

# Synthesis



# The Institutionalisation of Evaluation in the Americas: A Synthesis

*Reinhard Stockmann and Wolfgang Meyer*

## INTRODUCTION

The countries of North, Central and South America covered in this volume show considerable differences with regard to their governmental and social systems, their cultures, traditions, languages and economic strength. Looking at the **USA** and **Canada**, which have gross national incomes per person (GNI) of US\$ 65,760 respectively US\$ 46,370, the spectrum ranges from the world's richest countries with functioning political systems, an active civil society and a highly developed education system to countries that are among the poorest on earth, such as **Bolivia** (GNI of US\$ 3,530) (World Bank, [2020](#)). Between these, there

---

R. Stockmann (✉) · W. Meyer  
Center for Evaluation (CEval), Saarbrücken, Germany  
e-mail: [r.stockmann@ceval.de](mailto:r.stockmann@ceval.de)

Department of Sociology, Saarland University, Saarbrücken, Germany

W. Meyer  
e-mail: [w.meyer@ceval.de](mailto:w.meyer@ceval.de); [w.meyer@mx.uni-saarland.de](mailto:w.meyer@mx.uni-saarland.de)

are many countries with emerging market status that have made enormous economic and social progress in recent decades and have developed their education systems efficiently. The case studies in this volume deal with countries ranging from territorially huge dimensions such as **Brazil** to small state entities such as **Costa Rica**.

Accordingly, it is not surprising that the institutionalisation of evaluation in the Americas shows very different forms and characteristics. This can be seen at first glance. What the similarities and differences between them are, is to be examined in detail below. This includes a systematic comparative analysis of the country case studies. First, the institutionalisation of evaluation in the political system is examined by using the analysis guideline on which this study is based. Since it is clear that the legal and organisational anchoring of evaluation does not automatically mean that evaluations are actually carried out in reality, this question will be analysed next. Among other things, in order to determine the degree of institutionalisation of evaluation in social systems, it will be examined whether civil society in any way uses evaluations for its own purposes, whether it acknowledges and discusses evaluation results or even knows the concept of evaluation at all, whether it is actively involved in evaluations, or even consciously demands from political and social actors that evaluations are being conducted.

Next, the supply side is addressed. In order to find out whether the demand for evaluation from politics and society can be met quantitatively and qualitatively, the professionalisation systems of the individual countries are being compared with each other.

In a further analytical step, the three subsystems examined will be analysed in terms of their interrelations. The aim is to identify clusters in terms of the way in which the political and social demand for evaluation is structured—and terms of the degree of professionalisation on the supply side—and to detect any linkage between these aspects.

Finally, an attempt is made to identify the driving forces for the observed developments and the challenges that the evaluation systems of the countries under study will have to face in future.

Methodically, the case studies are assessed with the given analysis grid as a cross-sectional structuring aid. This entails a number of difficulties. Firstly, information is not available for all elements of the analysis grid in all case studies. Secondly, the terminology is not always used in a consistent way—for example, the term ‘evaluation’ is sometimes used in a very broad sense, encompassing measurement, audit, monitoring and impact

assessment. This is mainly due to the fact that the concept of evaluation is interpreted differently in each country. The authors of the case studies are well aware of these differences, but the actors in the individual countries do not observe these terminological subtleties.

Despite this slight fuzziness, which can also be seen when other terms are used, a meaningful overview of the institutionalisation of evaluation in the Americas can be given.

## THE INSTITUTIONALISATION OF EVALUATION IN THE POLITICAL SYSTEM

A legal framework is necessary so that evaluation can take place in the political system which includes governments with their ministries and authorities (executive) and parliament (legislative). The extent to which it is present in the Americas should first be examined (Table 16.1, columns 2 & 3).

It is remarkable that in some states, evaluation is even anchored in the constitution:

- In **Colombia**, “the function of designing and organising the evaluation systems of management and results achieved by public investment in the country” has been delegated to the National Planning Department (DNP) by the constitution of 1991 (p. 203).
- The **Costa Rican** constitution of 1949 (Article 11) already includes that “public administration shall be subject to a procedure of performance evaluation and accountability” (p. 247).
- In **Ecuador**, evaluation has been established in the constitution since 2008 as “principle of public administration” (Article 227, p. 274).
- The **Mexican** Constitution (Article 134) also provides that: “all programmes and budgets of the agencies and entities that involve economic resources of the federal government must be evaluated” (p. 301).

In other countries, evaluation is regulated by national laws that do not explicitly refer to evaluation, but to institutions to which evaluation has been assigned as a task. These can be ministries of planning (e.g. **Colombia, Chile, Costa Rica, Ecuador**) or ministries of finance (e.g. **Chile, Peru**) and Treasury Boards (**Canada**). Only in **Mexico** a national

**Table 16.1** Legislative institutionalisation of evaluation and evaluation use (own development)

	Institutionalisation of evaluation					Use of evaluation				
Country	National laws and strategies <sup>a</sup>	National decrees <sup>b</sup>	Organisational embedding <sup>c</sup>	Role of evaluation in audit office <sup>d</sup>	Mean	Role of parliament <sup>e</sup>	Sectoral spread <sup>f</sup>	Scope of evaluation practice <sup>g</sup>	Mean	Overall Mean
Argentina	0	1	0.5	0	<b>0.38</b>	0	0	0	<b>0.00</b>	<b>0.21</b>
Bolivia	0	1	0	0	<b>0.25</b>	0	0	0	<b>0.00</b>	<b>0.14</b>
Brazil	0	1	0.5	0	<b>0.38</b>	0	0	0	<b>0.00</b>	<b>0.21</b>
Canada	1	1	1	1	<b>1.00</b>	0.5	1	1	<b>0.83</b>	<b>0.93</b>
Chile	1	1	0.5	0	<b>0.63</b>	0.5	0.5	1	<b>0.67</b>	<b>0.64</b>
Colombia	1	1	0.5	1	<b>0.88</b>	0	0.5	0.5	<b>0.33</b>	<b>0.64</b>
Costa Rica	1	1	0.5	0	<b>0.63</b>	0	0	0	<b>0.00</b>	<b>0.36</b>
Ecuador	1	1	0.5	0	<b>0.63</b>	0	0	0	<b>0.00</b>	<b>0.36</b>
Mexico	1	1	1	0	<b>0.75</b>	0.5	1	1	<b>0.83</b>	<b>0.79</b>
Peru	1	1	0.5	0	<b>0.63</b>	0	1	0.5	<b>0.50</b>	<b>0.57</b>
USA	1	1	1	1	<b>1.00</b>	1	1	1	<b>1.00</b>	<b>1.00</b>

**Table 16.1** (continued)

<sup>a</sup>Here, only those laws were included that have a comprehensive national validity. As national laws, only such laws are considered, which were passed by the parliament (existing = 1, not existing = 0).

<sup>b</sup>As national decrees, all regulations are considered which are not laws (as described under 1), but refer to the whole nation (i.e. not only to particular policy fields) (existing = 1, not existing = 0).

<sup>c</sup>Hereby is meant, the existence of specific evaluation units in government institutions (such as ministries) and/or the existence of independent stand-alone bodies of evaluation (existing in both categories = 1, existing in one category = 0.5, existing in no category = 0).

<sup>d</sup>National audit offices carry out not only performance audits (which are limited to the evaluation of goal achievement (effectively) and/or efficiency), but also carry out evaluations with a broader focus (yes = 1, no = 0).

<sup>e</sup>Two criteria were evaluated: A) Parliament has an evaluation unit at its disposal and commissions evaluations. B) Parliament regularly takes note of and discusses evaluation results. (If both criteria are met = 1; one criterion = 0.5; no criterion = 0).

<sup>f</sup>Degree of spread across sectors, number of sectors (policy fields) mentioned in the case studies in which comprehensive evaluations are carried out (above 7 = 1; between 6 and 5 = 0.5; below 4 = 0).

<sup>g</sup>Intensity and frequency by which evaluations are carried out. Qualitative assessments according to case studies. Ratings in the text in brackets.

institution for evaluation has been founded by law, the ‘Council for Evaluation of Social Development Policy’ (CONEVAL). In the **USA**, the Government Accountability Office assumes the task of evaluating national programmes covering all sectors. In addition, the Government Performance and Results Act (GPRA), enacted by law in 1993, ensures that all major federal agencies are required to conduct evaluations (p. 359).

In all American countries, evaluation is not only regulated by national laws, but also by national decrees. Only three countries limit regulation of evaluation on decrees and did not endorse laws or acts (Table 16.1 column 3):

- There are no “laws establishing an integral evaluation system” (p. 43) nor is there a national “regulatory framework” in **Argentina** (p. 62), but instead there exists a decree that instructs the *Consejo Nacional de Coordinación de Políticas Sociales* to prepare an annual national M&E plan (p. 42). However, there are “no policies or consolidated strategies regarding evaluation” (p. 45) and, therefore, “evaluation as a policy depends more on the will of individual actors than on the institutional conviction of the state” (p. 62).
- In **Bolivia**, various national decrees emphasise the need to conduct evaluations (pp. 66ff.).
- In **Brazil** “there is still no national law establishing parameters and guidelines for the evaluation of policies in an institutional manner” (p. 100). A Policy Monitoring and Evaluation Council (CMAP) was approved by decree (pp. 102ff.), an inter-ministerial arrangement, that calls upon other institutions for evaluations.

For the Americas, it can be stated that evaluation, to a large extent, is regulated by legislation. Without exception, all countries do have national legislation and/or national decrees. In some countries, evaluation even has constitutional status. The legal regulations often do not refer specifically to evaluation itself, but to institutions (e.g. ministries or specially created authorities) that are entrusted with the task of evaluation.

For the organisational institutionalisation of evaluation (Table 16.1, column 4) you find both in the Americas, decentralised and centralised solutions and also combinations of both. In the following countries, evaluation is mainly organised by the ministries themselves:



- Because there is neither any regulatory framework nor a national evaluation policy and no agency or unit “in charge of explicitly planning, implementing and coordinating evaluation of programmes and policies” (p. 41), in **Argentina**, the institutionalisation of evaluation is described “as low, fragmented and isolated” (p. 54). Evaluations are conducted only occasionally by ministries and government agencies.
- In **Colombia**, the National Management and Evaluation Results System (SINERGIA) delegates the implementation of evaluations to the sectors. There exists “no specific evaluation regulation for each of these sectors” (p. 208). The DNP assumes only a coordinating role (p. 208).
- No centralised M&E system has been institutionalised in **Peru** either. Instead, the practices of the Ministry on Economic and Financial Affairs with regard to results-based management are adopted by other ministries (p. 327). In order to meet the legal requirements, “different sectors HAVE created instances in charge of monitoring and evaluation” (p. 329).

In five other countries, however, there are national institutions which, with a few exceptions (such as **Mexico**), are not specifically set up as organisations to conduct evaluations, but are merely entrusted with evaluations in addition to other fields of activity. These are often planning ministries:

- In **Costa Rica** where MIDEPLAN has been entrusted by law since 1974 “to monitor and evaluate the progress and objectives of the PND” (National Development Plan) (p. 247). The Ministry is now responsible for implementing the national evaluation plan (pp. 243f.), which is based on a national policy on evaluation (p. 246).
- In **Ecuador**, the National Planning and Development Secretary (SENPLADES) was assigned by law to plan and conduct evaluations (p. 273).<sup>1</sup>
- In **Chile**, even two national evaluation systems can be found. They are part of the Ministry of Social and Family Development (MDSF),

<sup>1</sup> SENPLADES was dissolved by decree in 2019. Some of the former planning and evaluation functions are now carried out by *Secretaría Técnica Planifica Ecuador*. For the consequences, see the **Ecuador** case study.

in the Undersecretary of Evaluation, and in the Ministry of Finance, in the Budget Office (DIPRES) (p. 172).

- Although a Federal Monitoring and Evaluation Council (CMAP) was created in **Brazil**, it has lost much of its political power in recent years due to personnel and government changes (pp. 103f.).
- **Mexico** is a special case insofar as CONEVAL was founded in 2005 as an autonomous, national and independent organisation with the task to assess the impact of programmes according to the “Performance Evaluation System” (SED) (p. 300). There are also other sectoral evaluation institutions such as the National Institute for Education Evaluation and the National Institute of Ecology and Climate Change.

**USA** and **Canada** represent a special case since central and decentralised organisational anchoring of evaluation go hand in hand:

- In the **USA**, evaluations were already being conducted in the 1950s and extensive organisational structures have been created for this purpose ever since. Within the executive branch, the key players in M&E are the Office of Management and Budget (OMB) and the individual federal agencies. While the OMB is mainly concerned with budget preparation, the individual federal departments and agencies directly collect the most federal performance data and manage most evaluations conducted by the government (cf. Mark & Pfeiffer, 2011, chapter 2.1). Even small agencies have some formal unit responsible for evaluations and are required to recruit appropriately qualified personnel trained in evaluation (pp. 359f.). In addition, the legislative branch has a great influence on the evaluation activities of the executive branch. And that is because “the U.S. Congress, through its legislative authorisation and appropriations committees in both the House and the Senate, plays a powerful role in overseeing and shaping executive branch policies and programmes” (cf. Mark & Pfeiffer, 2011, chapter 2.2). The Government Accountability Office is used by parliament for this purpose.
- In **Canada**, the ‘Policy of Results’ decrees that the Treasury Board of Canada Secretariat (TBS) is responsible “to require departments to undertake specific evaluations and participate in centrally-led

evaluations”.<sup>2</sup> TBS, which has similar functions as the US-OMB, dominates the national evaluation process by issuing rules and standards and monitoring the national evaluation system. In contrast to this, evaluations are conducted by evaluation units within individual government departments and agencies that measure the performance of their programmes and policies, making the Canadian model a poster child of the centralised policy making and oversight plus decentralised implementation (pp. 147f.).

In summary, it can be said that the organisational anchoring of evaluation is centralised in most countries. Only in some countries it is left to the ministries in a decentralised manner. In one case (**Mexico**), an independent national institution for evaluation was founded (CONEVAL), and in the **USA** and **Canada**, the federal agencies and departments have established extensive evaluation structures and also have strong independent external evaluation institutions.

Another feature of the political institutionalisation of evaluation is the role that audit courts play in the evaluation system of a country (Table 16.1, column 5). This assumes a special weight for evaluation in the political system when audit courts also conduct evaluations for their own purposes. But almost no audit court in American countries is doing so. Instead, they are conducting audits that focus exclusively on financial control. This is also the case in **Brazil**, where the audit court is supposed to conduct evaluations in the future (when the law is passed by Congress) (p. 113).

Currently, the only exceptions to this rule are **Canada**, **Colombia** and the **USA**. The **Canadian** Office of the Auditor General (OAG) conducts various forms of audits (e.g. legal compliance, financial and performance audits) as well as evaluations and “as conductor, coach, focusing on training and development and setting standards”, the OAG has an enormous influence on evaluation practice in the various ministries (p. 147). In **Colombia**, the *Contraloría General de la República* is evaluating the various organisations and bodies of the state, applying the principles of economy, efficiency, effectiveness, equity and environmental sustainability. The US-Government Accountability Office (GAO)<sup>3</sup> is responsible

<sup>2</sup> See: [www.tbs-sct.gc.ca](http://www.tbs-sct.gc.ca).

<sup>3</sup> “The GAO was established as the General Accounting Office in 1921, focusing on financial oversight of the executive branch. In 2004, it was renamed the General

for conducting evaluations in the **USA**. 3,300 employees are producing around 1,000 reports per year and testimonies and it “provides the legislative branch with a major means of assessing the effectiveness of government programmes. [...] the GAO’s activities now encompass a broad range programme evaluation, policy analysis, and legal opinions about all government programmes and agencies” (Mark & Pfeiffer, 2011, chapter 2.2).

While the fields dealt with so far are describing the legal and organisational anchoring of evaluation, the following topics are primarily intended to examine the sectoral breadth, intensity and scope of evaluations being conducted as well as their use. Of particular interest here is the role of parliaments. Their role relates not only to their law-making power, that is their ability to pass evaluation laws or enshrine countries to instigate or even directly commission evaluations. On top of this, parliaments can be important users of evaluation results. The situation in the individual countries examined is shown below (Table 16.1, column 7).<sup>4</sup>

In many parliaments, evaluation plays no role or only a marginal one.

- In **Argentina**, the relationship between M&E and parliament is described as ‘minimal’: “There is no regular practice regarding the report of the results generated by the evaluation of programmes and public policies” (p. 48).
- In **Bolivia**, there are no regulations on whether and how the parliament should deal with evaluation (p. 71).
- In **Brazil**, the parliament also does not show any interest in evaluation results. The reason given is that: “We still have legislative houses very focused on political game, using the role of laws for political bargaining” (p. 110f.).
- The **Colombia** report says, “the parliamentary structures sometimes use the results of evaluations for their political work” (p. 210).
- The **Costa Rican** report states that “some progress has been made by the legislative assembly to include evaluations and the use of results by parliamentary structures” (p. 249). However, this is rarely

Accountability Office to reflect a broader understanding of its mission” (Mark & Pfeiffer, 2011).

<sup>4</sup> Here, once again it should be pointed out that the 0.5 point score was only awarded if parliament actually performed regularly and not just occasionally takes note and discusses evaluation results (see remark under Table 1).

the case and legislative evaluation is not an institutionalised practice (p. 249).

- The **Ecuador** report states “that politicians tend to not even consider evaluation as an opportunity to learn and improve public policies” (p. 286).
- The **Peru** report stresses that the subject evaluation “is absent in the Congress of the Republic” (p. 351).

This is totally different in the **USA** and **Canada**. In Latin America, only the parliaments of **Chile** and **Mexico** seem to deal with evaluation results in more detail.

- In the **USA**, the parliament is not only the addressee of the reports of the Federal Agencies, but also has direct effects on the M&E activities of the agencies “both directly through its internal committee hearing process and indirectly through requirements that it imposes on the executive branch” (Mark & Pfeiffer, 2011, chapter 2.0). In addition, parliament can entrust the GAO with conducting evaluations: “Any member of Congress can request an assessment of governmental programmes, with priority given to Congressional mandates and leadership requests” (Mark & Pfeiffer, 2011, chapter 2.2.b).<sup>5</sup>
- In **Canada**, parliament is informed about the results of evaluations in various ways. The individual departments are under obligation to submit an annual Departmental Results Report (DRR) to parliament, which also takes into account evaluation results. This is meant to inform “Parliamentarians and Canadians of the actual performance and results achieved against plans by government organisations” (p. 154). However, the Canadian case study also states: “Nonetheless, the national parliament and parliament in provinces and territories do not play a significant role in the demand or the supply of evaluations” (p. 154).
- In **Chile**, all evaluation reports are submitted to the Congress. Also: “Congress can complement or suggest new evaluations according

<sup>5</sup> Mark and Pfeiffer (2011) give an excellent summary on the M&E system in the **United States** Government.

to the priorities of the legislative branch or according to requirements that arise during the budget discussion process” (p. 177). The evaluation results represent a “fundamental input” for the annual budgeting process (p. 182).

- In **Mexico**, CONEVAL delivers an evaluation report to the House of Representatives that is used by the parliament for its budget decisions (p. 304).

Overall, the interest of parliaments in the evaluation topic seems to be rather low. They hardly appear as users of evaluation, and as a rule, they do not commission evaluations. Exceptions are **Chile** and **Mexico**, where evaluation results are integrated into budgeting processes. Only the **USA** stands out, as the GAO is not only a strong institution that can be entrusted with the implementation of evaluations, but also has a direct impact on the evaluations of the Federal Agencies and uses evaluation results from both procedures for political decision-making.

Evaluation results can only be used for decision-making if they are of an appropriate quality. Therefore, the next step is to assess evaluation practice and the use of evaluations.

According to evaluation practice, there are certain differences between policy fields and sectors. As shown in Table 16.2, evaluations are often performed in Education, Health and Social Development and some countries even have their own evaluation units in such ministries and authorities. Evaluations are rarely conducted in the sectors Defense, Information and Communication, Transport, Aid, Foreign Policy and Home (internal) affairs.

Even in the **USA**, where evaluations are conducted in 20 out of 22 sectors, which are subject to the GPRA, there is a high concentration on a handful of sectors of financial resources spent on evaluation. The Department of Health and Human Services accounts for about one-third of all evaluation resources. This is followed by the Department of Veteran Affairs (about 20%) and the US-Agency for International Development (approximately 12%). This means that 3 of the 20 ministries have spent around 65% of the financial resources on evaluation available via the GPRA.

Across all countries, a wide range of sectors is covered by evaluations, but as can be seen from Table 16.3—with the exception of the two North American countries—only two Latin American countries cover a wider range (seven or more sectors) (**Mexico**, **Peru**). In five countries

**Table 16.2** Sectors where evaluations are conducted<sup>6</sup>(cf. Table 16.1, column 8) (own development)

<i>Sector</i>	<i>Number of entries</i>
Education	11
Health	10
Social (policy) Development/Social Welfare	10
Agriculture	6
Labour	6
Environment	6
Science, Technology, Innovation	5
Economy, Trade and Tourism Industry	5
Housing and Urban Planning	4
Justice	4
Culture	4
Defence	2
Information, Communication	2
Transport	2
Aid	2
Foreign Policy	1
Home (internal) Affairs	1

(**Argentina, Bolivia, Brazil, Costa Rica, Ecuador**), the sectoral spread is significantly lower.

The case studies say little about the scope and frequency with which evaluations are carried out in the individual sectors, whether internal or external evaluations and whether process or impact evaluations are mainly used, who does make primarily use of the results as well as how the quality of evaluations is rated and how it is assured.

In the following, however, an attempt will be made to extract the scope, frequency and intensity of evaluation practice from the case studies beyond the sectoral spread (Table 16.1, column 9).

- **Canada** and the **USA** have the longest tradition of evaluation and have institutionalised wide-ranging regulations that not only ensure a wide sectoral breadth but also a high intensity.
- In the **USA**, (1) 22 of the non-defence departments that were subject to the GPRA in 2017 spent US \$561 million, an increase

<sup>6</sup> It is an enumeration of the sectors mentioned in the case studies.

**Table 16.3** Degree of spread across sectors where evaluations are conducted on a regular basis<sup>7</sup> (cf. Table 16.1, column 8) (own development)

<i>High (1)</i>	<i>Medium (0.5)</i>	<i>Low (0)</i>
Canada	Chile	Argentina
Mexico	Colombia	Bolivia
Peru		Brazil
USA		Costa Rica
		Ecuador

*Note* (1) = spread in seven sectors or more

(0.5) = spread in five to six sectors

(0) = spread in four sectors or less

of more than 40% compared to 2010 (\$394 million). The GAO conducts approximately 1,000 evaluations per year (Mark & Pfeiffer, 2011, chapter 2.2b). Since the end of the 2000s, predominantly external evaluations have been conducted. These include both, process and outcome evaluations. In addition to think tanks and universities, independent consultants play a central role for the implementation of external evaluations (Mark & Pfeiffer, 2011, chapter 2.0). Besides people responsible for the programme as well as programme managers, the parliament is one of the most important users of evaluation results. According to the authors of the case study, nothing can be said about the use of evaluation results in the individual sectors, since there simply does not exist any universal evaluation control and monitoring agency in the USA (p. 366).

- **Canada** (1) is also intensively engaged in evaluation. During the Fiscal Year 2016–2017, 142 evaluations were conducted by department and agency evaluation units and US \$42.6 million were spent (including salaries of 437 fulltime equivalent employees which represent 78% of this amount as well as external contracts). These figures show that evaluations in Canada are mainly carried out internally (p. 150): “by agents who are members of the administration in charge of the implementation of programme under evaluation”

<sup>7</sup> Ministries with their own evaluation unit or objective of regular evaluations conducted by external evaluation institutions such as planning ministries or special evaluation institutions (such as CONEVAL).



whereby it is assumed “that they are granted latitude to do professional work”. Often cost/benefit analyses are carried out because the accountability aspect has been in the foreground since the Federal Accountability Act (2006) at the latest (p. 153). Even though “internal evaluators are not subject to formal professional requirements”, a competency framework was issued 2019 by the Treasury Board of Canada Secretariat (p. 151). Also, each Department has a senior Departmental Evaluation Committee that oversees the practice and use of evaluation (p. 151). The main users of evaluation are senior and programme managers, both being focused on instrumental use (p. 155).

- CONEVAL (**Mexico**) (1) which is responsible for evaluating social development, the only organisation in America specifically established to conduct evaluations at national level, had a budget of around 500 million pesos (US\$ 23.5 million) in 2019, which represents 0.4% of the national budget. This sum includes the salaries of the 204 staff members of CONEVAL. Between 2007 and 2013, about 1,700 external evaluations of federal budgetary programmes were conducted which means 243 per year. Of these, 98% were design, consistency and outcome evaluations, performance specific and others, and only 2% were process and impact evaluations (p. 306). The case studies do not report anything about the quality of the evaluations and quality assurance measures.
- As stated above, **Chile** (1) actually has two M&E systems. The MDSF states that a total of 532 projects from more than 10 ministries were evaluated ex-ante during the period from 2012 to 2018. Otherwise, programmes are mainly being monitored (p. 175). The DIPRES names four types of evaluation, which evaluate the consistency of programme goals, management aspects, use of resources and results with different emphasis. Impact evaluations make up only a small proportion of these. Between 1997 and 2017, DIPRES has evaluated a total of 533 programmes of 50 institutions. Measured in terms of the “evaluable budget”, 14% were evaluated between 2011 and 2017 (p. 179). The main purpose of the M&E system is “contributing to the quality of public spending” (p. 182). For this reason, “the results of these systems are effectively integrated into the budget cycle, improving efficiency in the allocation and use of public resources (...)” (p. 182). Accordingly, besides

programme managers, one of the main user groups are parliamentary decision-makers who use the results for budgetary allocation (p. 184). The case studies do not report anything about the quality of the evaluations and quality assurance measures.

- In **Colombia** (0.5), as well as in all other Latin American countries, “the number of projects evaluated each year is not significant compared to those implemented” (p. 205). Between 1997 and 2020, 271 evaluations were conducted (arithmetically about 12 per year), but with figures fluctuating very strongly each year (p. 206), so there is no trend observable. Even though evaluations were conducted in 20 of 24 sectors, four sectors only were accounting for 40% of all evaluations (Health and Social Protection, Presidency, Social Inclusion, Science and Technology). The majority of evaluations are conducted as external evaluations, two-thirds of which are outcome and impact evaluations (pp. 208f.), which is a very high proportion when compared with the other countries. The professional groups, which regularly use evaluation results, are mainly programme and policy managers, but hardly decision-makers (p. 210). The case studies do not report anything about the quality of the evaluations and quality assurance measures.
- In **Peru** (0.5), the respective ministries are representing the guiding entities of each sector which are developing an annual evaluation programme. A distinction is made between Evaluations of Budget Design and Execution (EDEP) and Impact Evaluations. Both evaluation types are performed externally (p. 332). Between 2008 and 2018, the Ministry of Economics and Financial Affairs (MEF) has commissioned 58 EDEPs (meaning about five per year), with a significantly decreasing frequency since 2013. Instead, since then, eight impact evaluations were carried out. The Ministry of Social Inclusion and Development has conducted a total of 32 evaluations (meaning about five per year) between 2012 and 2018 and the Ministry of Women and Vulnerable Populations a total of nine evaluations between 2015 and 2017 (meaning three per year). Further evaluations were conducted by the Ministry of Agriculture and Irrigation, the Ministry of Production and the Ministry of Housing. The case studies do not report anything about the quality of the evaluations and quality assurance measures.
- The **Brazilian** (0) evaluation system is described as “insufficient and unsatisfactory” (p. 114). The evaluations of government

programmes are characterised “by dispersion and discontinuity”. This is explained with two dominant features of the country’s government planning: a preference for the formulation of plans and the development of programmes on the one hand, and on the other one with the “high negligence in the stages of monitoring and evaluation processes, results, and impacts” (p. 114). Regarding the use of evaluation, the following statement is interesting “that evaluation in Brazil has more of a diagnosis role, rather than being used to improve or correct public policy” (p. 117). The Court of Auditors has characterised the government’s M&E activities as extremely deficient and proposed “that the legislature can contribute to this improvement, making clear instruments for institutionalising these processes” (p. 117).

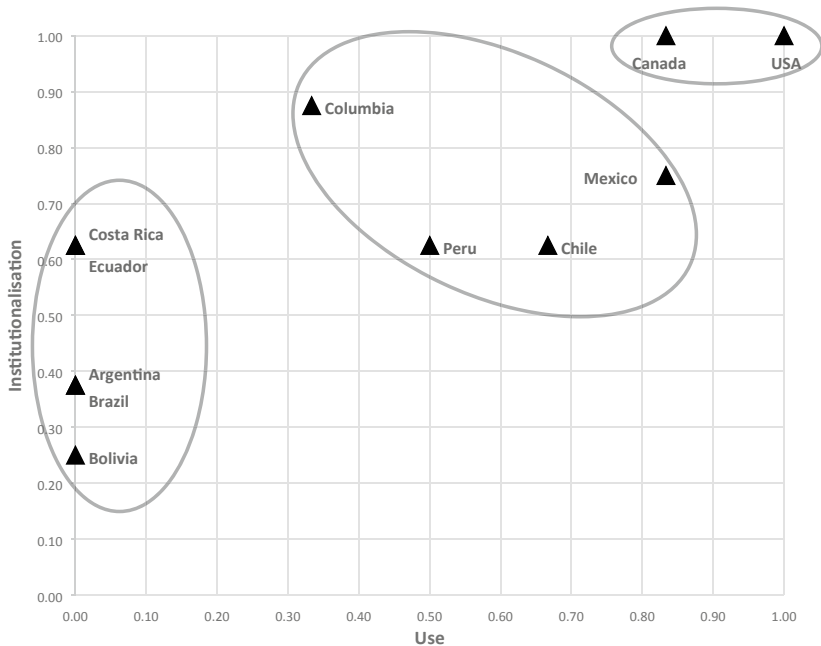
- For the period from 2015 to 2018, **Costa Rica’s** (0) national evaluation plan has foreseen 15 evaluations of projects and national programmes of different institutions, but not all of them were realised. In addition, a majority was donor funded. Mainly, they were outsourced or conducted jointly with MIDEPLAN and the ministries concerned (p. 252). Since evaluation is largely perceived as a control by those concerned, evaluation has a bad image and results are often not used (p. 253). In addition, only very small budgets are made available for evaluation, so that the implementation of evaluations still depends very much on the funds provided by international donors, whose evaluations “are very focused on accountability and goals achievement” (p. 253).
- **Argentina** (0) also lacks a developed evaluation culture and a regulatory framework (p. 54). Therefore, only few evaluations are being conducted. Most ministries do practise evaluation. If they do, the main focus is on the budgetary and operational perspective (p. 53). A “strategic framework integrating monitoring and evaluation functions into the daily management of the public administration” is missing (p. 53).
- In **Bolivia** (0), only very few evaluations are conducted, and when they are conducted, it is often because of donor requirements (p. 75). The quality of evaluation is affected because there are hardly any qualified personnel (p. 76). When evaluations are carried out, the results are mainly used by the programme management and those responsible for the programme. Political decision-makers use evaluation results—if at all—for their own benefit, depending on whether

evaluations support a political situation or not, but they never use them to make technical decisions” (p. 75).

- In **Ecuador** (0), mainly due to a lack of financial resources and qualified personnel, evaluations often only take place when international donors finance them (p. 276). That is why the frequency to which policies and central government programmes are evaluated is erratic (p. 277). Moreover, most of the so-called evaluations are monitoring activities (p. 277). Since there does not exist any well-established culture of evaluation up to now (p. 282), the few existing evaluation results are only used by a limited number of people, especially programme management, staff responsible for administration and academia. Politicians, on the other hand “in general do not know about evaluations nor use evaluations” (p. 284).

If one compares the degree of legal and organisational institutionalisation, as expressed in criteria (1) existence of national laws, (2) national degrees, (3) organisational embedding in governmental organisations, (4) role of evaluation in audit courts in Table 16.1, with the criteria that say something about the use of evaluation like (5) role of evaluation in parliament, (6) degree of sectoral spread and (7) scope of evaluation practice—then one gets the following result shown in Fig. 16.1.

- (1) With its decades of evaluation practice, the **USA** is clearly proving to be outstanding and is setting a standard in this area. The USA has not only a legal framework that over time has been repeatedly modernised and, with the GPRA a detailed set of rules and regulations, but also interdepartmental implementation structures. These include both the executive and the legislative branches. This is a special feature when comparing the American countries. The federal departments and agencies are collecting performance data and are conducting evaluations. On the legislative side, the GAO acts independently of the executive branch and within the area of responsibility of the parliament, which, compared to all other American countries, plays a prominent role regarding the commissioning and use of evaluations. In addition, the GAO “has long been a strong proponent of efforts to build and strengthen executive branch M&E capabilities and practices” (cf. Mark & Pfeiffer, 2011, chapter 2.2.b).



**Fig. 16.1** Institutionalisation and evaluation use in the Americas (own development)

The GAO emphasises that “the GPRA Modernisation Act can offer opportunities to help make tough choices in setting priorities as well as reforming programmes and management practices to better link resources to results” (U.S. GAO, 2011, according to Mark & Pfeiffer, 2011, chapter 2.2.b). The recommendations of the GAO are not formally binding, but it is reported that over 80 per cent of them are implemented (same here).

**Canada** can also be considered a very positive benchmark. The evaluation is regulated comprehensively by law, there are central standards for its implementation and there is the TBS oversight, which has similar functions as the US-OMB. TBS is a management board for government functions including evaluation, performance measurement and internal audit (besides a number of other tasks

such as financial-, human resource- and IT-management as well as budget operations).

Like the **USA** with its GAO, **Canada**, with its OAG, has a Supreme Audit Office that conducts a variety of audits, including those that can be described as evaluations. Both hire evaluators and both operate under audit standards. In organisational terms, the **Canadian** model represents a combination of central rule-making and oversight with decentralised delegated implementation, mostly (in contrary to the **USA**) through internal evaluations. The role of the parliament is differently developed in the **USA** and **Canada**. While the U.S. GAO receives assignments for evaluation directly from parliament, the Canadian parliament gets only so-called DRR, which may have used evaluation results. In **Canada**, parliaments do not play an important role in the demand or supply of evaluations. As in the **USA**, the range of evaluations covers all areas that evaluate regularly and with great intensity.

- (2) A group of four countries (**Mexico, Chile, Colombia** and **Peru**) is characterised by a medium to high degree of institutionalisation as well as a medium to high degree of implementation and utilisation. All four countries have national evaluation laws and national decrees. In all cases, the evaluation is embedded in the organisations.
- (3) The third group of countries is formed by those, which have only a medium (**Costa Rica** and **Ecuador**) or low (**Argentina, Brazil** and **Bolivia**) level of institutionalisation. In these countries, evaluation is hardly ever used.

It should also be noted that for the use of evaluations it makes no difference whether the evaluation function is anchored centrally or decentrally. In three countries (**Argentina, Colombia** and **Peru**), there is a decentralised setting, while in five others (**Brazil, Chile, Costa Rica, Ecuador, Mexico**) the evaluation function is centralised and regulated by institutions commissioned to conduct evaluations. This demonstrates that both forms of organisation are qualified to carry out the implementation of predetermined legal regulations.

Figure 16.1 can still be interpreted to mean that the legal and organisational anchoring of evaluation in a country is the best guarantee that evaluation will be used for political practice. Apart from the **USA**, however, this statement applies only to a limited extent. Countries such

as **Costa Rica** and **Ecuador** also have at least a middle degree of legal and organisational institutionalisation, but the use of evaluation is very low in **Costa Rica** and almost non-existent in **Ecuador**. This indicates that although the degree of legal anchoring of evaluation is related to the extent to which evaluation is used in political practice, the institutionalisation of evaluation in no way guarantees its use. Furthermore, the other way around it can be stated that there is no country in which there is no legal framework, and yet evaluations are carried out to large extent. On the contrary, in those three countries (**Argentina**, **Bolivia**, **Brazil**) where there is only a marginal legal framework to be found, hardly any evaluations take place.

The audit offices do not seem to have any influence on the contexts described, since they do not perform any evaluation tasks except for **Canada**, **Colombia** and the **USA**.

Further observations can be summarised as follows:

- In most countries, evaluation is used for accountability purposes. This is certainly due to the fact that many evaluation activities are linked to budgeting and financial control processes and are used by planning ministries, among others, for monitoring purposes (see **Argentina**, p. 50, 62; **Brazil**, p. 108; **Chile**, p. 175; **Colombia**, p. 210; **Costa Rica**, p. 251; **Ecuador**, p. 279, 283; **Peru**, p. 330; **Canada**, p. 151). In some countries, there are more monitoring than evaluation activities involved (see **Argentina**, p. 51; **Bolivia**, pp. 67f.; **Brazil**, pp. 114ff.; **Costa Rica**, p. 251; **Ecuador**, p. 282).
- In addition, despite a glossary of binding terms, evaluation was not used in a uniform manner in all case studies. It is not always obvious whether the term also includes audits and performance measurements, or whether it simply refers to monitoring activities, thereby overestimating the number of evaluations actually conducted. Whereas the authors of the case studies are aware of these differences, the actors in the individual countries usually do not observe these terminological subtleties. Little can be said about the type of evaluations conducted. In **Argentina** (p. 49) and **Canada** (pp. 150f.), evaluations are primarily conducted internally, while external evaluations are dominating in **Chile** (p. 183), **Colombia** (p. 208), **Peru** (p. 332) and the **USA** (p. 361).
- In those cases when evaluations are financed primarily by international organisations, these are mostly external evaluations that also

focus on accountability (see **Costa Rica** p. 251, **Ecuador** pp. 277f.). The proportional share of impact evaluations of all evaluations in the individual countries cannot be determined, but it appears that it is very low in most countries (except **Colombia, USA**).

- Moreover, only little is known about the quality of the evaluations carried out and institutionalised quality assurance mechanisms. Just a few reports provide information on this. In **Bolivia**, the “weak evaluation training offers”, the “high mobility of personnel” and the lack of regulations for quality (p. 76) are blamed for the poor quality of the few evaluations conducted. In **Costa Rica**, too, there is criticism that there exist “no minimum standards or requirements that define the quality of evaluations or the skills of evaluators” (p. 253). In **Ecuador**, the complaints are that: “there ARE a lot of poor quality evaluations” (p. 285), because there is not any quality assurance.

### THE INSTITUTIONALISATION OF EVALUATION IN THE SOCIAL SYSTEM

Talking about social systems in America is special because of the role of civil societies. The scientific concept of civil society is derived from the **USA** and, therefore, some authors call it ‘the American model’: civil society is seen as a product of liberalism, especially as a “system of free and unfettered associability, pluralism, and largely unregulated interest group or nongovernmental organisation (NGO) activity” (Wiarda, 2018, p. 3). Moreover, if one calls civil society ‘the American model’, he or she has a powerful, self-assured, independent and well-organised community in mind that understands society as a task for citizens. It includes a certain scepticism about the capability of the state and sees the civil society in charge for social development.

This U.S. model is different to the European welfare state with a partnership approach for civil society. This is derived from a long battle between state and civil society, formed as social movements to address democratic changes in society. ‘The Swedish model’ of civil society is nowadays more a ‘family approach’ and an alternative draft for the concept of civil society (Micheletti, 2017). This model spread over Europe and reached also the South—with a particular connotation due to poor state performances (Busso & De Luigi, 2019).



In Portugal and Spain, the civil society became a driving force for democratisation and bringing both countries to the European Union (Fernandes, 2015). ‘The Spanish model’ reveals that civil society is “rather dependent for its effects on the wider political context” and “contemporary democratic transitions are now traditionally associated” with economic disruptions and scepticism about the (new) political institutions—and “in turn, such conditions cast a dark cloud over the long-term viability of both civil society and democracy” (Encarnación, 2001, pp. 74, 79). In general, this model seems to be more appropriate to describe the situation in the Global South, particular in Latin America with its alteration between democracy and dictatorship in many countries during the last century.

All these ideas and approaches somehow merged in Latin America, building a very unique form of civil society and the discussion on social movements in Latin America is even going “beyond civil society”: the civil society agenda in Latin America is nowadays “a hegemonic though contested set of normative and prescriptive assumptions about citizen participation” and because of these confrontations, “unruly political action by ‘uncivic’ society inherently threatens democracy” (Alvarez et al., 2017, p. 2). As a result, social movements in Latin America are further divided and the weak civil society is more critically examined than in North America.

These are important framework conditions if one looks at the institutionalisation of evaluation in civil society in the Americas. On the one hand side, there are strong civil societies including rich associations, foundations and companies (**Brazil, Canada, Mexico and USA**), and on the other hand, some states can be characterised by weak, fragmented and less independent civil societies (**Bolivia, Costa Rica, Ecuador and Peru**). However, the effects of size and strength on the institutionalised use of evaluation in civil societies are poor. Almost everywhere, evaluation is seen as an activity of the public administration and the initiative for including the civil society in evaluation debates is coming from the political system and not from civil society. However, there are no formal participatory institutions including evaluation into the dialogue between public administration and civil society like, for instance, referendums or consultations (Table 16.4, column 2). In some countries, there are pilot projects or isolated cases (**Brazil, Canada, Chile, Ecuador**), but there is no pressure from civil society for getting more involved by using evaluations for decision-making.

In **Chile**, “Civil society has timidly begun to envision evaluation as a public good for advocacy or accountability” (Chile, p. 185). In **Colombia**, “the space for community participation has been expanded and the use of this practice is growing” but “this practice seems not to be fully institutionalised” (Colombia, p. 213). However, there is some evidence for evaluations and their results being for supporting referendums. In **Costa Rica**, evaluation results are not used for political decision-making and they are just reaching an initial level of awareness. “It is therefore no surprise that sectors outside the public administration do not even have knowledge of the evaluation processes, their implications and benefits, which leads to drastically reduced levels of use of evaluations and results” (Costa Rica, p. 254). In **Argentina**, it is almost the same situation: “the presence of civil society organisations in instances close to the monitoring and evaluation of government actions has been minimal and marginal” (Argentina, p. 56). In **Bolivia** and **Peru**, there is also no visible use of evaluation in the dialogue between state and civil society.

Not surprisingly, the situation is a bit different in the four big countries with a better developed civil society. Already in 2011, Mark and Pfeiffer (2011, p. 14) stated for the **USA**: “A key feature of the U.S. M&E environment is the vigorous role played by nongovernmental organisations both in overseeing and assessing government policies and programmes—including funding the work of experts who are monitoring and evaluating government programmes—and also in working to influence government decision making in both the executive branch and Congress. Some of the most influential evaluations of government programmes in the United States have been the product of collaboration among government agencies, private philanthropic foundations, and research institutions”. One indicator for the engagement of civil society in evaluation is the membership in AEA: more than 40 non-profit organisations are involved in the work of this Voluntary Organisation for Professional Evaluation (VOPE) and there is a non-profit and foundation topical interest group in American Evaluation Association (AEA). Moreover, there are certain initiatives of evaluators to engage civil society for evaluation (USA, p. 367).

However, there is remarkably no evidence for using evaluation as an instrument to control or supervise state activities systematically (Table 16.4, column 2). Although civil society is much more engaged in evaluation than in many other countries, one cannot say that the U.S. civil society is pushing the state towards evaluation or significantly contributes to the development of evaluations of state programmes. Evaluation use

is merely limited on own purposes to improve the performance of civil society organisations and to learn for their own activities. Civil society does not use evaluation as a tool for societal development—at least not in a systematic and institutionalised way.

The situation in **Brazil** and **Canada** seems to be rather similar, as stated here for **Brazil**: “Although some government monitoring and evaluation initiatives are being strengthened by civil society action, evidence indicates that civil society has had more of a role in training and supporting the public sector than in demanding evaluations through bureaucratic mechanisms of participation, for example” (Brazil, p. 124). Almost the same for **Canada**: “While the general public (or ‘civil society’) in Canada would generally not initiate a formalised systematic evaluation, there are many formal and informal elements that bring evaluation into the public sphere” (Canada, p. 155f.). Although there is a strong and self-assured civil society, a widespread understanding and a certain culture of evaluation, civil society is in none of these three countries a driving force for the development, implementation and use of evaluation.

The situation is slightly different in **Mexico**. While the civil society is also a strong and highly diversified group of ‘multi-sector actors on different areas’ with more than 40,000 registered CSOs, being far away from a homogenous and somehow synchronised unity, the critical perspective on state activities seems to be a bit better installed than in the other Latin American countries. There are think tanks like *México Evalúa* (<https://www.mexicoevalua.org/>) who are regularly monitoring and evaluating state activities for supporting evidence-based policy. The National Council for Evaluation CONEVAL honours “good practice in the use of monitoring and evaluation results” and some civil society organisations received prizes. These measures seem to have some effects—with a rather limited scope: “although there are CSOs that have already internalised the use of evaluations as standard practice, and these results are visible, for the vast majority there is no documented information” (Mexico, p. 312). Hence, civil society is also no driving force for evaluation development in **Mexico**, even if there are some encouraging institutions to support and improve evaluation use.

Particularly in these four countries with a strong civil society—**Brazil**, **Canada**, **Mexico** and the **USA**—a lot of activities promoting evaluation may be able to attract public interest (Table 16.4, column 4). Although there are some examples for the dissemination of evaluation results in **Brazil**, there is “no concrete evidence that evaluations are perceived by

the general public as an important tool for policy improvement” (Brazil, p. 127). For **Canada**, the authors recognise that “the media are either not tuned into the evaluation culture or they lack the resources to use evaluation information” (Canada, p. 158). Moreover, there seems to be a strong difference in using the word ‘evaluation’ in media: while in English, media prefers to use ‘review’, the French media chooses evaluation but link it merely to environmental topics. In **Mexico**, CONEVAL organises conferences and other public activities, invites media to these events and encourages them to report on evaluation activities. However, although there is “a greater degree of knowledge about the existence of evaluation ... this has not yet permeated into the society” (Mexico, p. 313). There is also no evidence for the **USA** that promotion activities by public or private actors had led to a deeper understanding or widespread public discussion of evaluation. To sum it up: public interest in evaluation is very low—even if huge public or private corporates are promoting it.

Not surprisingly, the situation is not better in countries without such strong promoters. In **Chile**, **Costa Rica** and **Ecuador**, no use of evaluation in the general public has been observed and discussions are limited to a small group of experts, merely programme managers. In **Peru**, most NGOs (about 100) grouped into the National Association of Research and Social Promotion Centers (ANC), aiming on “improvement of the quality of democracy in the country”. Although ANC was built 30 years ago, it “does not demand neither the use of evaluations, nor spaces for the debate, analysis and reflection on public policy evaluation” (Peru, p. 340). In **Bolivia**, a survey revealed the focus of evaluation use on own purposes (learning for programme development) in civil society—nobody reported about the use of evaluation results for public discussions (Bolivia, p. 79). In **Argentina**, “evaluation experiences carried out within the national public administration scope are not usually socialised in a broad manner” (Argentina, p. 56) and this is seen as one reason for the lack of interest to use evaluations in public debates. Civil society, particular the foundations, “show a minimal and marginal interest for evaluation and concentrate their action on citizen training on state control bodies (their functions, reporting mechanisms, the information they produce, etc.), the formation of electoral observatories, construction of indexes regarding the perception of corruption, etc.” (Argentina, p. 56).

All these struggles for bringing evaluation into the public are obviously not very successful. All countries report a lack of knowledge about evaluation in the broader public (Table 16.4, column 3). Moreover, one has to

recognise that activities for making evaluation more popular are coming from public institutions and not from civil society. The poor knowledge about evaluation is not caused by missing transparency of state institutions and there are no indices for pressure of civil society to be more transparent according to evaluation reports.

In three countries, the access to evaluation reports is regulated by laws or decrees and the availability of evaluation reports (Table 16.4, column 5) is a legal right. For the **Argentina** case, there is a decree regarding the access to public information that aims at strengthening the relationship between state and civil society by laying all official documents open—including evaluation results (Decree No. 1172; Argentina, p. 56). In **Canada**, the public institutions are even forced by law “to support uptake by increasing user-friendliness of evaluation information” and “the policy went further than previous versions by requiring the posting of evaluation summaries, in addition to the full report – the intent being to support uptake by increasing user-friendliness of evaluation information” (Canada, p. 156). However, this does not change knowledge and awareness of these reports: “while federal evaluation reports are freely available on government websites, they are not the subject of active promotion or of a communication strategy” (Canada, p. 157). In **Chile**, “there is the possibility for any citizen to request the reports from the agency, and this must be accepted by the institution as mandated by Law 20,285 on Transparency of Public Function and access to information of the State Administration” (Chile, p. 188). Evaluation reports are frequently published by ministries and other public institutions on their website. This may also be true for most other countries, as well as the following description: “dissemination of results is carried out almost exclusively through the publication of reports on web pages without a greater focus on the dissemination strategies of results, which limits their scope and potential use” (Chile, p. 187).

In all other countries, there are no obligations for making evaluation reports available although there is some evidence that it is usual practice of public administrations to publish evaluation reports on their website (or at least a summary of the whole report). As stated in **Costa Rica**—but probably universally valid—“the fact that the reports are public is not a guarantee of knowledge and use within the institutions themselves, nor within sectors outside these institutions” (Costa Rica, p. 257). Even if national authorities—as mentioned above for CONEVAL in **Mexico**—are actively promoting the dissemination of evaluation results, there is no

evidence for an increased knowledge about and use of evaluation reports in public discourse. In **Brazil** and the **USA**, the authors mentioned a limited practice to publish evaluation results. “With respect to the policy evaluation practices [in **Brazil**], 38% of the organisations affirmed that they have practices of disclosure for the external audience” (p. 125). The situation seems to be comparable in the **USA** (p. 368), where “any evaluation services not directly contracted by the federal government are not required to make data, results, or recommendations accessible to the public”, and especially at the local level, it is “difficult for those interested to find and use the results of evaluation”. Although there seems to be a remarkable positive development on the national level for providing evaluation results, there are at least some difficulties on local level in large countries. However, even if there is some evidence for a widespread practice of public supply, there might be other hindering factors that prevent access. For instance, the authors of the Columbian and Peruvian case express some doubts about knowledge and use of public accessible sources. This might be a general phenomenon although there is hardly any evidence to approve this assumption.

As mentioned above, there is no pressure from civil society by demanding evaluations (Table 16.4, column 6). While there are some exceptions, the overwhelming majority of civil society actors does not recognise evaluations as an instrument for holding governments and public authorities responsible—at least in a regular and systematic way. In some cases, it seems to be the other way round: civil society organisations are doing evaluations because of the pressure of funders (which are in many cases public authorities). The survey in **Bolivia**, for instance, shows that 16% of civil society organisations are doing evaluations due to demand of funders (Bolivia, p. 75). Although there is no evidence on this for the other countries, one can assume a reasonable minority of civil society organisations being not convinced by evaluation as a management instrument and being forced by third parties like national authorities or international funders to do evaluations (see also the results from the Brazilian survey, Brazil, pp. 124f. and from a study in Ecuador, pp. 286f.). The dissemination of evaluation in civil society is not an endogenous process deriving from internal discussions or public debates. The pressure is coming from outside and it is a slowly evolving process to convince civil society organisations of the advantages of evaluation primarily for their own programmes and projects. Remarkably, the huge differences in

civil society concepts between North and South America do not lead to comparable differences in development or forms of institutionalisation.

If evaluation is used in civil society, the main argument seems to be learning for their own purposes and not to control state activities. This may be best illustrated by the evaluation philosophy of the **USA**—based on the Bill and Melinda Gates Foundation, the world’s largest private philanthropical foundation with almost 50 billion dollars capital and about 1,500 employed people: “Achieving our ambitious goals requires rigorous evaluation so we and our partners can continually improve how we carry out our work”.<sup>8</sup> One can find comparable statements in other US foundations like the Ford or the Rockefeller Foundation. While still “many non-profits view evaluation activities as a resource drain on already limited funding and a distraction from their core mission” and “fail to see the added value evaluation can provide” (USA, p. 366), the number of private organisations using evaluation as a tool at least in programme management for own purposes is rising. However, at least ten years ago, “non-profits that receive federal funding engage in considerably more programme evaluation than non-profits primarily funded by foundations or local governments” (USA, p. 366). To sum it up: even in the **USA** as the country with the best organised civil society, a huge non-profit sector with a citizens engagement lens and the longest tradition in evaluation, still many private organisations have to be convinced of the use of evaluation for their own purposes—and if evaluation is used, it is primarily used for the purpose of organisational learning and not for societal development, pushing authorities to more or more specific use of evaluation as an instrument of decision-making.

The situation is not very different in **Bolivia** as one of the countries with a much weaker civil society and less developed non-profit sector. As the results of a survey show, about half of the responding organisations emphasise “learning good practices” as the aim of evaluations and 82% address evaluation results primarily to programme management (Bolivia, p. 75). In general, demand for evaluation from private organisations “is weak, in a context where evaluation culture is absent and expertise is not recognised as a public value” (Bolivia, p. 81). A study in **Ecuador** indicated “that although private enterprises implement programmes, they usually do not demand evaluations. When private enterprises conduct

<sup>8</sup> See: <https://www.gatesfoundation.org/How-We-Work/General-Information/Evaluation-Policy>.

evaluations, they tend to be process evaluations” (Ecuador, p. 289). The situation seems to be very similar in **Peru**: “the concept of evaluation use is restricted, both in the public and private sector, to programme management, reaching external aid sources and, exceptionally, programme or project stakeholders, but not the general public” (Peru, p. 338). In **Costa Rica**, there is no information available on evaluations carried out by private organisations (Costa Rica, p. 254).

This poor situation is not limited to these small countries, but can also be found in the larger ones with a more improved third sector and a stronger civil society. In **Argentina**, evaluations done by civil society “usually focus on the performance of the programme or project, assessing results and drawing lessons that allow the improvement of the NGO’s or funding agency’s general programming” (Argentina, p. 55). The authors also highlight that most of these performance assessments are done by inner members of the implementing organisation. The 2018 civil society census in **Brazil** came to a comparable result: “From the organisations that evaluate their programmes, 74.4% said that they perform evaluations with their internal team during the execution of the programme” (Brazil, p. 125). Most of these evaluations are also performance assessments and impact evaluations by external professionals seem to be still exceptions. In **Canada**, there is only a small demand for evaluation coming from private companies or civil society and the authors assume that in many cases these evaluations are only done for symbolic use (Canada, p. 157). In **Chile**, the authors recognised a certain “demand for developing social impact evaluations regarding Corporate Social Responsibility strategies—or similar—of private companies” (Chile, p. 186) that already influenced the evaluation market.

The most positive developments can be found in **Colombia** and **Mexico**. Due to state regulations on project control in **Colombia**, social organisations are under state observation and a certain pressure to evaluate their projects. As a result, “demand for evaluation in Colombia by social organisations has improved” (Colombia, p. 216). The signing of the Peace agreement in 2018 gave another push for social programmes and this may help to overcome the non-satisfying situation twenty years ago when evaluation was an infrequent, even exceptional activity. **Mexico** faces also some serious changes in the Federal Public Administration and this has opened space for more public debates. This will probably have a certain impact on evaluation culture and the evaluation practice in civil society. However, the current situation is still improvable: “Although



there are CSOs that have already internalised the use of evaluations as standard practice, and these results are visible, for the vast majority there is no documented information” (Mexico, p. 312).

As a conclusion, the institutionalisation in the Americas’ civil societies can be described as follows:

- In general, there is not much impact on the development of evaluation caused by activities and initiatives of civil society and its organisations.
- Although there are strong differences in size, strength, power, membership, philosophy, understanding and acceptance of civil societies between the countries, the variation in performance and institutionalisation of evaluation are low.
- There are four countries with a more powerful and improved civil society (**Brazil, Canada, Mexico** and the **USA**) and this results in a slightly improved use of evaluation both for developing societies in cooperation with (or addition to) state authorities and for own purposes. However, the differences are less shaped than expected.
- Even in the most advanced countries, one cannot find any kind of formal institutionalisation of evaluation developed by civil society itself. The development of evaluation institutions is clearly state-driven.
- This is especially true for the only institutionalisation of evaluation in the social system: in most countries, evaluation reports both from public and private organisations are available due to usual practice. In three countries (**Argentina, Canada** and **Chile**), this practice bases on acts and decrees endorsed by national authorities, giving citizens a right for access to public documents including evaluation reports. No comparable and independent institutions developed solely by civil society can be found in any American country.
- Evaluation is not used by civil society for holding the government responsible or for accountability reasons. There is no public discussion about evaluation and the tool is only known in a small circle of experts, merely from programme management. If civil society organisations use evaluation, they do it for organisational learning reasons, primarily in form of performance assessments done by internal staff. External impact evaluations are still rare exceptions, even in the **USA**.

**Table 16.4** Institutionalisation of evaluation in the social system (own development)<sup>a</sup>

	Institutionalised use of evaluations	Public perception and discussion of evaluation			Civil societies' demand of evaluations	
Country	General use of eval. in CS	Knowledge about evaluation	Public discussion/ media	Availability of reports	Civic demand	Mean
Argentina	0	0	0	1	0.5	0.30
Bolivia	0	0	0	n/a	0	0.00
Brazil	0.5	0	0	0.5	0.5	0.30
Canada	0.5	0	0	1	0.5	0.40
Chile	0	0	0	1	0.5	0.30
Colombia	0	0	0.5	1	0.5	0.40
Costa Rica	0	0	0	1	0	0.20
Ecuador	0	0	0	n/a	0	0.00
Mexico	0.5	0	0.5	1	0.5	0.50
Peru	0	0	0	1	0	0.20
USA	0.5	0	0	0.5	0.5	0.30

<sup>a</sup> 1 = high spread, 0.5 = medium spread, 0 = low spread, *n/a* = *not answered*.

INSTITUTIONALISATION OF EVALUATION  
IN THE SYSTEM OF PROFESSIONS

America is the birthplace of professional evaluations, particularly in the northern part of the continent. Already in the 1970s, the first courses, Master's and PhD programmes on evaluation (Table 16.5, column 2) had been offered in the **USA**. Several authors proofed the development of academic trainings at U.S. universities over time (LaVelle & Donaldson, 2016), and today, there are 50 evaluation-specific master's degrees, 35 certificate programmes and 40 doctoral programmes focused on preparing future evaluators (USA, p. 372). Hence, it is not a steady growing of offers but a highly fluctuating come-and-go of new courses and programmes (Meyer, 2016, p. 109). These offers cover a broad variety of different topics and there are no generalised curricula how to teach evaluation.

Nevertheless, the broad offer of academic training in evaluation at U.S. universities may have influenced the development of programmes on the whole continent. The ‘pull’ factor of U.S. universities—especially the most respected private ones like Harvard or Stanford—is traditionally very high in the Americas, although evaluation is primarily established in less-popular state universities like Western Michigan University in Kalamazoo, Michigan, or Claremont Graduate University in Claremont, California. Nowadays, one can find a broad spectrum of academic training offers, master and PhD programmes almost in all American countries. The strongest offers can be found in:

- **Argentina:** three Masters in Public Policy Evaluation, several minor subjects, programmes and single courses in particular fields (e.g. labour market, participatory evaluation, social policy, territorial development and public sector evaluation). There are even initiatives to foster the institutionalisation of evaluation in the academic sector of the country (Argentina, pp. 59f.).
- **Brazil:** three Masters in Public Administration and Public Policy Evaluation, several single and short time courses, “which is almost nothing taking into account the M&E field growth” and the size of the country (Brazil, p. 132).
- **Canada:** several universities offer trainings in evaluation at the undergraduate, master and doctoral levels both in the French-speaking province Quebec (one master and one doctoral degree program) and other parts of Canada (four universities offer graduate diplomas in evaluation and the University of Waterloo has a Master of Health Evaluation). Many universities offer specialised evaluation courses in various disciplines like criminology, education, psychology, public administration and public health (Canada, p. 159).
- **Colombia:** two Masters in Environmental respectively Educational Evaluation and 24 specialisations in evaluation as part of other study programmes. In total, 118 postgraduate programmes include at least one course related to evaluation (Colombia, p. 219).
- **Costa Rica:** The Master programme at the University of Costa Rica (UCR) is one of the oldest and well-known ones in the Americas and together with the Centre for Research and Training in Public Administration the UCR is offering a broad variety of academic training programmes for different clients (merely from public administration). This offer is not limited to Costa Rica but has also some

impact in the region. Moreover, several disciplines such as economic policy or social work and institutions like the Centre for Economic Policy for Sustainable Development have incorporated at least one evaluation component, almost always expressed in one subject at the undergraduate level (Costa Rica, p. 259).

In the other countries (**Bolivia, Chile, Ecuador, Mexico** and **Peru**), no master programmes or other forms of extended academic evaluation courses have been mentioned by the authors. In **Bolivia**, for example, “there are no formal academic offers aimed at professionalising monitoring and evaluation (M&E) in the country. Rather, there are specialised M&E courses within other disciplines” (Bolivia, p. 82). In **Chile**, “there is a wide training offer” but “it is mostly a diploma offer, not including an offer to obtain a Master’s or Doctor’s degree” (Chile, p. 191). In **Ecuador**, “there are some sporadic and isolated courses that are offered in programmes from other professional fields” (Ecuador, p. 290). In **Mexico**, 17 postgraduate programmes and 14 diploma or training programmes have been traced “that offer in their curricula a significant content of subjects, in their professorships, oriented towards the training of evaluators” (Mexico, p. 314). In **Peru**, there are six postgraduate programmes and 13 training programmes including Monitoring and/or Evaluation topics, “offered by Peruvian and international academic institutions in face-to-face and on-line learning mode” (Peru, p. 342). All these offers do not form a systematic academic programme on evaluation (not even for a minor or specialisation subject), but stay on the level of isolated single courses.

All countries report about additional non-academic offers and one has to mention here that obviously online-offers like webinars or e-learning platforms are playing an increasingly important role. Such offers are not primarily coming from country actors like administrations, consultancies, universities or national VOPEs but from international evaluation organisations and experts like, for instance, international organisations of development cooperation (German Institute for Development Evaluation/Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval), Inter-American Development Bank Group (IDB), International Fund for Agricultural Development (IFAD), United States Agency for International Development (USAID), World Bank) or initiatives of global or regional VOPEs (EvalPartners, Red de Seguimiento, Evaluación y Sistematización de Latinoamérica y el Caribe (ReLAC)). Besides these

increasing and heterogenous (from both the perspective of issues raised and content quality) online-offers, there are several other forms like, for instance, summer schools (mentioned in **Brazil, Canada** and the **USA**) or short time courses offered by specialised training centres sometimes linked to universities (mentioned in **Brazil, Canada, Costa Rica** and the **USA**). In general, the non-academic courses appear more often and more frequently in those countries with strong academic offers and there is no evidence for non-academic offers being able to substitute or replace academic ones.

Non-academic trainings are often offered by VOPEs, mostly linked to their annual meetings. In general, VOPEs are the most important institution for communication about evaluation among members as well as with stakeholders for the evaluation communities. In difference to training offers, the variation of the organisational degree is extreme, in size, influence and reputation. While some countries do not have any VOPE at all or only very small and merely informal groups (**Bolivia, Chile, Colombia, Costa Rica** and **Ecuador**), there are three dominating organisations in the Americas: the AEA with about 7,300 members in **USA**, the **Brazilian** Network for Monitoring and Evaluation (RBMA) with approximately 8,500 members and the **Canadian** Evaluation Society (CES) with more than 1,800 members (Table 16.5, column 4).

CES was the first VOPE in the world, founded in 1981, and is a professional association with paying members. CES is still a forerider for many reasons (e.g. it is the first evaluation society that created a professional designation credentialing programme in 2009) and therefore attracts not only Canadians. For instance, there is a cooperation with the European Evaluation Society and one can become a member of both organisations with reduced fees.

AEA was incorporated a few years later than CES and soon became the most important evaluation association in the world. As the CES, it is a professional association, including members from more than 80 countries spread all over the world. For many reasons, AEA is somehow the regional evaluation association and not only limited to the **USA** and topics derived from the States. The annual AEA conferences are important events for exchange and meetings, offering an opportunity to get together with the most famous (U.S.) evaluators.

RBMA is an open network that attracts about 8,500 people, registered as members, without a special membership fee, to get open access to the platform (including webinars, job postings and other material). It

includes merely consultants and academics and aims at strengthening the professional work of evaluators in **Brazil** (Brazil, pp. 131f.). It is the only network in Portuguese language, and therefore, it is somewhat isolated in the Americas.

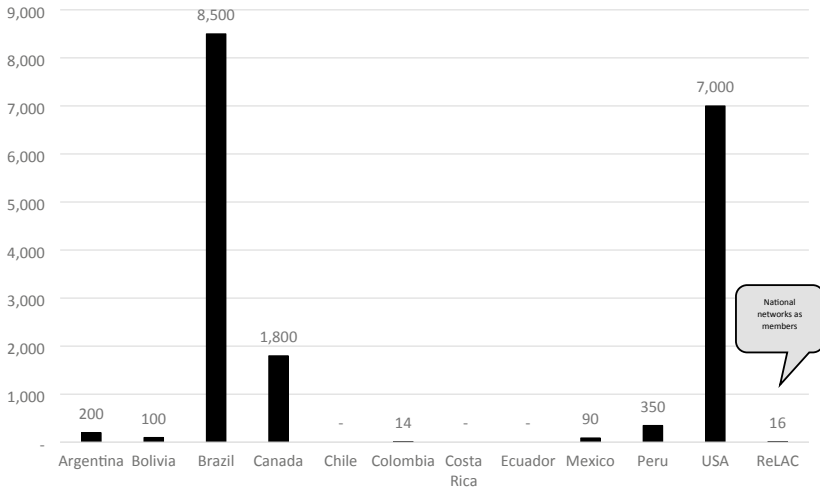
Besides the anglophone AEA and the lusophone RBMA, there is also a transnational Spanish-speaking umbrella network **ReLAC**<sup>9</sup> with 16 members representing the national networks of the region and about 2,200 visitors of the website. This open access network evolved from PREVAL (Programme for the Strengthening of Regional Evaluation Capacity of the Projects for rural poverty reduction in Latin America and the Caribbean), a programme supported by IFAD and the first trial to get Spanish-speaking evaluators together. ReLAC was also supported by internationals to provide its services, especially the ReLAC conferences.

The strengths of these transnational networks may explain the weakness of the national VOPEs in Latin America. Even the larger VOPEs in **Argentina** (Red EvaluAR) and **Mexico** (ACEVAL) are rather small compared to the size of these countries (200 resp. 90 members). REDMEBOL in **Bolivia** is an open network that reaches more than 100 people (Bolivia, p. 85). Some other national VOPEs do not even exceed 50 members or no information was available (Fig. 16.2).

In some countries (**Colombia**, **Ecuador**, **Peru**), the small evaluation community even split up into more than one network. These networks are more small working groups and not comparable with the huge associations in North America. Even RBMA and ReLAC are not as powerful and professionally managed as AEA or CES, merely because of the organisational form, the membership fees and—as a result—the budget of the organisations. However, there is one thing all VOPEs have in common: they are the focal point of communication about evaluation and this communication is channelled through a broad variety of media.

One of this media are scientific journals (Table 16.5, column 3) and VOPEs are important supporters for evaluation centred journals. Especially, the American Journal of Evaluation and the New Directions for Evaluation, both sponsored by AEA and distributed to all its members, and the Canadian Journal of Programme Evaluation, the only bilingual English/French Evaluation Journal, published by CES, belong to the most popular and respected evaluation journals in the world. RBMA is

<sup>9</sup> More information at: <https://www.relac.net/>.



**Fig. 16.2** Members in American VOPEs (own development)

supporting *Revista Avaliação em Foco* and revitalised *Brasileira de Monitoramento e Avaliação*, formerly edited by the Ministry of Social Development (until 2015), now as *Revista Brasileira de Avaliação* (RBAVAL). Both in **Brazil** (Brazil, p. 135) and in the **USA** (USA, p. 373), there are further scientific journals published by universities or (scientific) foundations with a certain link to evaluation topics. In **Colombia**, there is no particular journal for evaluation but the central register allows to search for evaluation articles in 275 A1 rated university journals. During the last decade, one-third of all articles in these journals are related to evaluation (Colombia, p. 220 f.), revealing the importance of evaluation at least for educational evaluation.

In all other countries (**Bolivia**, **Costa Rica**, **Ecuador**, **Mexico** and **Peru**), there are no journals or even newsletters available for communicating primarily on evaluation; only **Argentina** and **Chile** mention public administration journals as an option to publish on evaluation (Argentina, p. 60; Chile, pp. 192f.). **Costa Rica** reports about the recent initiative to establish a regional forum magazine on evaluation of development policies and some trials to embed evaluation in the broader spectrum of publications on Costa Rica's public administration (Costa Rica, p. 260). However, there is an astonishing lack of Spanish evaluation journals

although many national authorities and universities are involved in evaluation practice and research. If one wants to publish about evaluation in Latin America, he or she has to do it in English for an international community or maybe in Portuguese. Taking into account that many people in public administration are not familiar with these languages and professional communication in Latin America is almost exclusive in Spanish, this is a notable deficit in professionalisation of evaluation.

The gap between North and South America becomes visible if the focus is set on standards, guidelines, code of contacts and other instruments to guarantee a certain quality of evaluations (Table 16.5, column 5). For both **AEA** and **CES**, evaluator competencies, quality assurance and ethical principles in evaluation are important topics from the very early beginning and a lot of members engage in these issues. However, the first and globally most influential initiative is not coming from the VOPEs but from an independent committee. In 1975, several professional associations were concerned about the quality of evaluation practice and therefore formed ‘The Joint Committee on Standards for Educational Evaluation (JCSEE)’ which published a first set of standards on programme evaluation in 1981. Since then, the committee updated the original standards and developed further standard systems for specific topics in educational evaluation.<sup>10</sup> The JCSEE standards were adopted from VOPEs in many countries (among them **Canada**), were used as a blueprint for the development of own standard systems (among others the ones from **ReLAC**) or at least inspired discussion on standards.

**CES** not only adopted the JCSEE standards but has become a standing member of JCSEE. CES members have to approve knowledge and have to adhere to the CES evaluation standards. Furthermore, CES developed Codes of Conducts in the 1990s, currently under revision. Finally, CES is the first VOPE providing a credentialing programme in 2010, and until 2019, 428 evaluators were granted the Credential Evaluator designation (Canada, p. 162). **AEA** uses AEA guiding principles for recommendations on professional ethical and practical conduct of evaluation. There are no endorsed standards for evaluation but AEA supports initiatives like CHESS, a Checklist for Evaluation-Specific Standards (USA, p. 374). In both, **Canada** and **USA**, there are also other institutions setting standards and guidelines for evaluation. In general, the recommendations are more binding in **Canada** than in **USA**. For example, the Treasury Board of Canada endorsed standards for evaluation practice and they are binding for all federal evaluators (Canada, p. 164).

<sup>10</sup> For more information, see: <https://evaluationstandards.org>.



Compared to the extensive debates in the North and the number of documents endorsed by the VOPEs and other committees in **Canada** and the **USA**, the stage of institutionalisation in Latin America is still very poor. Hence, there is one remarkable exception: most recently, the transnational VOPE **ReLAC** approved the first standard system in Latin America. Supported by the German government and its regional project FOCEVAL (*Fomento de Capacidades en Evaluación*), **ReLAC** brought together representatives from almost all Latin American VOPEs and managed an intensive transnational exchange on evaluation standards. This was the first overarching initiative to develop a shared evaluation standards system and **ReLAC** endorsed its standards in 2016. Its success can be seen by the fact that some of the larger VOPEs (**Argentina**, **Chile** and **Mexico**) recommend to use the **ReLAC** standards instead of developing its own standards system. There is only one VOPE about to develop a national standard system (**Brazil**), but most countries do neither have their own national standard system nor officially endorsed the **ReLAC** standards yet (**Bolivia**, **Colombia**, **Costa Rica** and **Peru**). Moreover, there are some doubts about the use of these standards in every day's evaluation practice. In general, quality assurance seemed to be not an important issue in the Latin American evaluation communities in the past and it is too early to say whether the **ReLAC** standards may be able to change this for the future.

To summarise the state of professionalisation, one has first to consider the strong differences in the continent. While the North American countries **Canada** and **USA** are global foreriders with a broad variety of academic study programmes, well-established and internationally high recommended journals, huge associations and a certain set of norms and standards for evaluation (in **Canada** even with a certain degree of obligation), many Latin American countries (especially **Bolivia** and **Ecuador**, but also **Chile** and **Mexico**) are still at a very early stage of professionalisation. **Brazil**—somehow isolated due to the language—developed less stable institutions but seems to be in many aspects (study programmes, organisation of communication on evaluation) in front of the Spanish-speaking countries. Among them, **Argentina** and **Colombia** developed more professional institutions like the other countries. However, compared to the size of the country, the evaluation networks are still very weak and the professionalisation seems to be limited to certain disciplines (public administration in **Argentina**, education in **Colombia**).

**Table 16.5** Professionalisation index (own development)

Country	Education	Communication	Organisation	Norms	Mean
Argentina	1	0.25	0.25	0.25	<b>0.44</b>
Bolivia	0.5	0	0.25	0	<b>0.19</b>
Brazil	1	1	0.25	0	<b>0.56</b>
Canada	1	1	1	1	<b>1.00</b>
Chile	0.5	0.25	0.25	0.25	<b>0.31</b>
Colombia	1	0.25	0.25	0	<b>0.38</b>
Costa Rica	1	0.25	0.25	0.25	<b>0.44</b>
Ecuador	0.5	0.25	0.25	0	<b>0.25</b>
Mexico	0.5	0	0.5	0.25	<b>0.31</b>
Peru	0.5	0	0.25	0	<b>0.19</b>
USA	1	1	1	0.75	<b>0.94</b>

**CPEI Professionalisation Scale 0–1**

Scale	I1 education: Study programmes	I2 communication: Focused exchange	I3 organisation: VOPE	I4 norms: General agreement
0	No offer available	No offer available	No VOPE existent	No rules existent
0.25	Only non-academic offers	Exchange in other discipline media	Open network without duties	Informal agreements
0.50	Only single academic courses	Exchange in open media	Formalised network	Self-commitment on internal rules
0.75	Minor subject courses	Exchange in regularly published media	Small formalised organisation	Endorsed general rules
1	Major subject courses	Exchange in academic journals	Large formalised organisation	Obligatory rules and certifications

The professionalisation index also reveals the strength of academic institutionalisation in opposite to the other aspects. This is the only area where Latin American countries are not lagging significantly behind North America. On the opposite, there is no Latin American country with a powerful association and communication structure.

## CONCLUSIONS: DRIVING FORCES AND CHALLENGES

At a first glance, there is one clear result if one compares the three systems under investigation here: while there is a certain degree of institutionalisation of evaluation in the political system and in the system of professions, the institutionalisation in the social system is very poor. Although there are huge differences in the strength, power and historical role of civil societies in the Americas, there are almost no variations between countries. Even if evaluation is used by civil society organisations, this is caused by pressure from state authorities and linked to funding principles and aiming for more transparency of public spending. Civil society organisations are held responsible for the implementation of public programmes and projects—and this is the reason why evaluation became a management tool for the non-profit sector. Nowadays, the use of evaluation for performance measurement and for learning about implementation management is widely spread—but impact evaluations or participatory approaches developed for the own purpose of the organisation are still rare exceptions. Furthermore, there are almost no contributions for the development of evaluation approaches from this side although several (US-American) evaluators are engaged in inclusive evaluations and developed innovative concepts used in the context of civil society programmes (e.g. Fetterman, King, Mertens).

There is almost no demand for evaluations to be commissioned by national authorities articulated by civil society representatives. Evaluation results—although easily available in most countries—are not used for holding public administration responsible or for the purpose of improving transparency and participation in the country. American civil societies do not recognise evaluation as an instrument for themselves, but merely see it as an instrument used by the state for accountability reasons. Many authors assume that linking evaluation to state control is a strong hindering factor for the acceptance of evaluation in civil society. If this is true, the interdependence between the political and social system led to a barrier for the diffusion of evaluation within the social system and—although there are strong differences between civil societies—it was not possible to overcome this barrier in any of the American countries.

The correlation between the institutionalisation of evaluation in the political system and in the system of professions is weak. However, one can identify four different clusters of countries with specific linkages between the systems (Fig. 16.3).

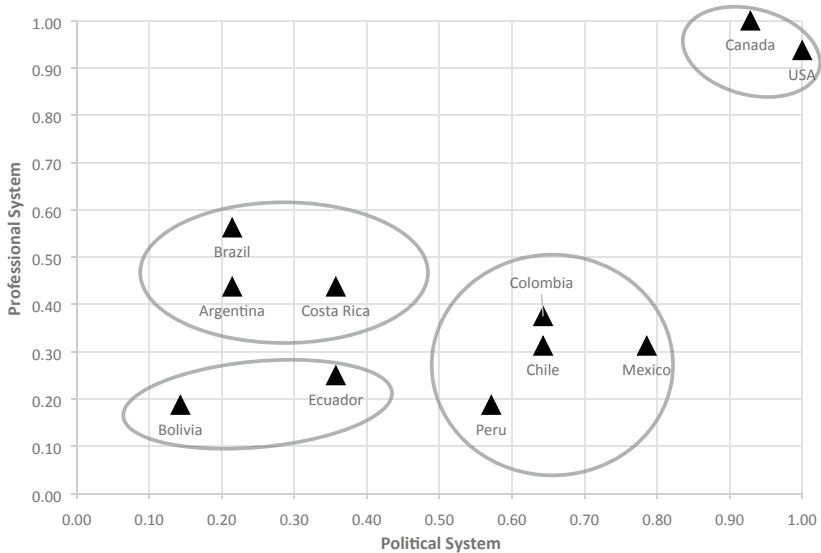


Fig. 16.3 Correlation political and profession system (own development)

The first group are those countries with a high degree of institutionalisation in both systems, namely **Canada** and the **USA**. Both countries belong to the pioneers in evaluation, from the perspective of implementing evaluation both in the political and in the profession system. There is a certain causal linkage between the systems in the 1970s/1980s (this will be described a bit more in detail later in this part), especially the developing and implementation of study programmes and associations were surely encouraged by developments in the political system and the increasing demand for evaluation there. However, there is a strong system dynamic taking over at least from the mid-1980s, leading to new challenges and a decoupling between both systems.

The second group is the opposite, including countries with a low degree of institutionalisation in both systems, composed of **Argentina**, **Brazil** and **Costa Rica**. Although there are huge variations, these countries also seem to have certain aspects in common. From the perspective of the political system, there is still a remarkable degree of legal institutionalisation, but the practical performance is rather poor—not many evaluations are commissioned although they are anchored in decrees,

laws and even in the constitution. This is somehow foiling the positive development in institutionalisation of evaluation at universities. There are well-developed study programmes and training courses for evaluation—but almost no market for young emerging evaluators. Moreover, there are no professional associations with a well-developed infrastructure for evaluation existing like in **Canada** or the **USA**. In general, the VOPEs are open networks with a poor degree of organisation, limiting the opportunities to be active for their members. Although both the **Argentinian** and the **Brazilian** VOPE are large networks, they are by far not as active and productive like AEA or CES in the North. To sum it up: these three countries are characterised by extremes—far going formal anchoring in academics and the legal system vs. poor performances in public administration and providing services for the growing evaluation community.

The third group including **Bolivia** and **Ecuador** possess only a very rudimentary evaluation community which is not well organised and also not fed by established training and study programmes. There is almost no evaluation practice both in the political and in the social system.

The final group is very interesting because it is ‘off the line’ from poor to good institutionalisation in both systems. Here, one finds four countries (**Chile**, **Colombia**, **Mexico** and **Peru**) with a high degree of institutionalisation in the political system but still a low level of institutionalisation in the national systems of professions. The main reason for the poor state of the professions is the lack of organisation. There is no strong evaluation association in Latin America although some of the networks comprise a great number of members: the Brazilian network is even larger than the American Evaluation Association and the members also have to pay some fees—but the organisational structure, the engagement in voluntary work for the organisation and the services offered are much higher in the **U.S.** organisation. Furthermore, the **Brazilian** network is probably more oriented towards lusophone countries than towards the rest of Latin America, while AEA and also the CES include a lot of Latin American members.

Due to the historical development, the anglophone association gained a special position in evaluation community and its work radiates beyond national borders. This may have had a negative effect on developing comparable institutions in the Spanish-speaking world and joins a weaker position of voluntary organisations in Latin America not limited to the field of evaluation. The main Spanish-speaking network is not bound to

nation states but established as a regional umbrella network, **ReLAC**. On behalf of organising conferences, for instance, ReLAC is the only important Latin American actor—although the national VOPEs are very active in organising such conferences. Hence, while RELAC also does not generate enough income and therefore is no strong professional organisation, the network depends on external support (e.g. the cooperation with the global organisation of evaluators IDEAS for the last ReLAC conference or the support of IFAD for founding such a network in the beginning of the new century).

Three of the four countries grouped here played an important role in establishing ReLAC as the nucleus of Latin American evaluation community. By looking at the five ReLAC conferences, the host countries were **Peru** (as the ‘birthplace’ of ReLAC), **Colombia** and **Mexico** (sublimated by **Costa Rica**) which is at least demonstrating that there are some active evaluators in these countries. However, only the VOPE in **Mexico** has a certain number of members and a bit more institutionalised organisation, although it is still a weak network compared to the strong organisations in **Canada** and the **USA**. Moreover, **Chile**, **Colombia**, **Mexico** and **Peru** belong to the countries with a certain high degree of institutionalisation in the political system, however, not on the same level as **Canada** and **USA**. The differences between these Latin American countries and for example Canada can be seen in the system of professions: while Canada built up a well-organised evaluation community, this does not happen—at least to the same degree—in these countries in South America.

While the variations in voluntary organisation between North and Latin America may be explained by cultural differences, the second important characteristic of this group of four countries is even more surprising. If one excludes North America and the two Latin American countries with almost no institutionalisation of evaluation (**Bolivia** and **Ecuador**), there is a negative correlation between the institutionalisation in the political system and in the system of professions ( $r = -0.75$ ). Countries like **Mexico**, **Colombia** or **Chile** with a high degree of institutionalisation in the political system are purely institutionalised in the system of professions, whereas countries like **Argentina**, **Brazil** and **Costa Rica** with a higher degree of professional institutions are less established in the political system.

The differences here are primarily caused by the academic system: while one can find well-established study programmes on evaluation in **Argentina**, **Brazil** and **Costa Rica**, especially **Mexico** and **Chile** do not

have any comparable offers. Furthermore, there are no evaluation journals or even non-academic publication media in Spanish language available in Latin America, while one can find both in Brazil and in Portuguese. In **Argentina** and **Costa Rica** (but also in **Chile** and **Colombia**), there are some journals in other disciplines that regularly publish evaluation results (e.g. public administration or education). These results indicate again the independence of both systems, revealing the surprising fact that political demand for evaluations and professional academic supply of evaluation competences are not linked to each other as expected.

### *Driving Forces*

Only the **USA** has an evaluation tradition that goes back to the nineteenth century (cf. Wollmann, 1984, p. 104). The start of ‘modern’ evaluation is considered to be part of the so-called ‘New Deal’ in the 1930s and 1940s (cf. Rist & Paliokas, 2002, p. 226). The midst of a boom in evaluation began in the early 1960s as a result of extensive social, educational, health and infrastructure programmes. The establishment of a Planning, Programming and Budgeting System in the mid-1960s was a very important impulse. This system, which was imported from the private sector, was a variant of the then popular systematic (management) approaches, which aimed to improve system effectiveness and efficiency and fund allocation by linking explicitly defined organisational objectives to output and input parameters (Stockmann & Meyer, 2013, p. 18). The administrative modernisation required new planning techniques and assessment methods. The necessary data could be obtained from evaluations, so that the demand for evaluation expertise rose sharply. One effect of this development was a fast increase of offers on study programmes and training courses in the early 1970s. No other country developed a comparable amount of professional trainings so early, merely targeting on different sectors in the process of modernisation of administration.

In the case of the North American neighbour **Canada**, which also ranks among the first-wave countries internationally (cf. Derlien, 1990, p. 147), “programme evaluation was born in the mid-seventies from a concern that the federal government had lost control over its finances” (p. 145). A major incentive for this was provided by the Auditor General, who demanded in a report of 1976 “that it was indispensable to set up performance indicators to measure effectiveness of programmes” (p. 146). The evaluation function was initially assigned to the Office of the Comptroller

General (cf. Segsworth, 2002, p. 175). The increasing amount of evaluation demand leads to establishing the first evaluation association in Canada in the 1980s.

In the other American countries, the development of evaluation started much later. **Colombia, Costa Rica, Chile and Mexico** were among the first countries in Latin America to begin setting up M&E systems and to establish the legal frameworks for them.

As in **Canada**, and already before it in the **USA**, evaluations in **Costa Rica** began to develop in the context of a changed administrative management. Here too, the decisive impulse came from financial management. As early as the mid-1950s, efforts were made in Costa Rica to improve financial control, but it took until the mid-1990s “thanks to a rethinking process of the economic model and the state structure that sought to modernise the Costa Rican State” (p. 242) to establish a National Evaluation System (SINE). Another outcome of this development was the establishment of a Master of Evaluation at the University of Costa Rica in 1995. This had been one of the first university study programmes in Latin America.

In **Colombia** and **Mexico**, it was also recognised in the 1990s that evaluation “plays an important role in the modernisation of the government apparatus and the impact of their investments” (Kliksberg & Rivera, 2007, p. 130, cited after Mexico, p. 300). In **Mexico**, a start was made in the mid-1990s to build up a “regulatory framework of the monitoring and evaluation system (M&E) aimed at achieving results” (p. 300). Also in **Colombia**, “following the New Public Management (NGP) and the Results-Oriented Management (GPOR) approaches”, a M&E System was implemented in the 1990s “to rationalise and make public action more effective, as well as to improve the efficiency and effectiveness of the state” (Colombia, p. 202).

**Chile** is also one of the frontrunners who started to establish an M&E system in the 1990s. After the recovery of democracy, a couple of initiatives were started “aiming to modernise public management” (p. 172). In the case study, Chile is called a pioneer “in the use of M&E systems in the framework of the budget cycle in the region and it has fully installed a budgeting for results approach” (p. 196).

At the beginning of the twenty-first century, other Latin American countries followed this trend and for the same reasons. In **Peru**, Budgeting for Results was implemented in 2007, which is described as a “milestone marking the development of evaluation in the country”



(p. 327), because it “introduced a change concerning budgeting, based upon an integrated vision of planning and budgeting and the articulation of actors and actions oriented towards the achievement of results” (Shack & Rivera, 2017, cited after Peru, p. 330).

The appearance of evaluation processes in **Ecuador** at the end of the 2000s “goes in hand with the institutionalisation of planning processes” (Ecuador, p. 272), as also in **Bolivia** (p. 66) and **Argentina** (p. 43), where the focus was on the evaluation of public investment programmes. In **Brazil**, “the creation of the first Public Policy Monitoring and Evaluation Council, the CMAP”, represents a milestone, partly as a result of a severe financial crisis at the end of the 2010s (Brazil, p. 99).

Overall, it can be stated that the development of evaluation had its start in the **USA**. After that, there were various waves of implementation: **Canada** followed in the 1970s and then, in the 1990s, **Colombia**, **Costa Rica**, **Chile** and **Mexico** formed the Latin American frontrunners group. **Peru**, **Ecuador**, **Bolivia**, **Argentina** and **Brazil** did not start using evaluations until the beginning of the twenty-first century or even until the 2010s.

In all countries, the decisive impulse to initiate an administrative modernisation oriented towards New Public Management concepts was coming from the fiscal authorities. In most cases, it was about improving national control and establishing budgeting for results systems. Therefore, evaluation was mainly used for accounting, for control purposes. Later, after financial control, other performance criteria such as goal attainment and outcome measurement were included, but impact evaluations are still the exception. Moreover, evaluation still suffers from this control image today, especially in civil societies with a certain distrust in government activities due to historical reasons. Therefore, in many cases, evaluation in Latin America is perceived as a threat rather than as an instrument for enlightenment. For this reason, hardly any country has a political dialogue based on transparent evidence from evaluations. Only project and programme managers still use evaluation as a learning tool to intervene in implementation processes and sometimes even for organisational learning.

It is true that some **international organisations** and **national donors** have strengthened this understanding of evaluation, as for decades they also have used evaluation primarily for accountability purposes. But it is undoubtedly their merit that evaluation has become increasingly widespread in Latin America. Especially, two international organisations

and a multi-donor network (the IDB, the World Bank and the CLEAR-initiative) have supported the development of M&E structures and actively practised evaluation capacity building, as the three institutional reports of **CLEAR LAC**, **IDB/The Office of Evaluation and Oversight (OVE)**, **IEG/World Bank** comprehensively demonstrate in this volume.

A particularly favourable constellation arose in **Mexico**, because **CLEAR LAC** benefited from the Mexican government's evaluation-friendly policy, which in turn benefited from CLEAR's activities, so that in the 2000s **Mexico** became an international reference for other Latin American countries (CLEAR LAC, pp. 395f.).

In **Costa Rica** (p. 243), the **IDB** has contributed significantly to the development of the national evaluation plan SINE. Since the mid-2000s, Costa Rica has been strongly supported by German development cooperation and university cooperation projects in the establishment of governmental M&E structures and in training and further training.

In **Brazil**, the World Bank, CLEAR, the IDB and the W. K. Kellogg Foundation have conducted evaluation capacity building with the result that "a generation of researchers and managers was formed based on international cooperation" (Brazil, p. 97). It is also mentioned in **Ecuador** that "International Development Agencies have been the main source of finance, planning and evaluation" (Ecuador, p. 271). In **Bolivia** (p. 78) and **Argentina** (p. 43), evaluations are usually only conducted if they are financed and promoted by national and international donors. The **Bolivia** report (p. 78) states: "the issue of evaluation is mostly maintained as a requirement of international cooperation", without creating a spillover effect. International organisations continue to play an important role especially for the evaluation latecomers.

The contributions in regard to the institutionalisation of evaluation are described as follows by the international organisations.

The **Independent Evaluation Group (IEG)** of the World Bank Group not only rightly points out that as a global player it is difficult to separate IEG's influence in the Americas from that of other actors when working in partnerships, but also emphasises that IEG itself was and is, of course, also exposed to influences (p. 425). Located in Washington in the USA, the United General Accounting Organisation, renamed the Government Accountability Office in 2004, the main audit body of the Congress, has influenced the evaluation understanding of World Bank and IEG. It is evident by looking at the original name *Government Accounting*

*Office* that the origins in the **USA**—as observed in **Canada** and in many **Latin American countries**—are in auditing, and that evaluation activities only developed further in the following decades from accounting to increased outcome and impact measuring. This is a stage of development that is still missing in most Latin American countries.

The **IEG** has not directly led projects and programmes on the development of M&E systems in Latin America, but rather indirectly through some advisory services and in its function as a role model for other Regional Multilateral Development Banks and International Agency Networks and Initiatives, by the development and provision of guidelines, policies, standards, manuals, analyses, textbooks and especially by evaluation capacity development. This includes the support of VOPEs, the establishment of two **CLEAR** Centers in the Americas (as a multi-donor initiative, including with the **IDB**) and the International Program for Development Evaluation Training (IPDET) in Canada, which moved to Switzerland in 2018. In 2020, **IEG** is establishing a new global partnership programme and multi-donor trust fund—the Global Evaluation Initiative (GEI)—in cooperation with global evaluation stakeholders (donors, technical partners and others). GEI’s partners will support Evaluation Capacity Development (ECD) in developing countries, fostering evidence-informed decision-making through enhanced M&E frameworks, capacity and use (p. 444).

The **Inter-American Development Bank Group (IDB)** has influenced the institutionalisation of evaluation in Latin America in two particular ways:

- (1) Through its operational work, given that many IDB-financed projects include the evaluation of specific project components. These evaluations generate evaluation capacity within counterpart agencies through learning by doing (IDB, pp. 409f.) as evaluations focus on borrower and beneficiary activities and are directly implemented by them. In addition, **IDB** finances projects that support national country systems related to statistics and to M&E. To date however, the focus has been on monitoring systems (IDB, p. 412) and less on evaluation.
- (2) Through the activities of its evaluation offices, including among others, capacity building and “the creation of a variety of evaluation tools and resources” by the self-evaluation office (IDB, p. 413); and the active dissemination by OVE of its evaluation work and

oversight studies, and its support of the **CLEAR** centres in **Mexico** and **Brazil** (IDB, p. 415).

One of them, **CLEAR LAC (Mexico)**, has presented its activities in this volume. The **CLEAR** centres generally provide training services, technical assistance and knowledge generation and services with a region-specific mix and locally differentiated sectoral and thematic foci. They typically engage with national and subnational government agencies and legislative bodies (p. 390). **CLEAR LAC** has been offering these services since 2010 and, among others, was “directly involved in the discussion on national evaluation policies and strategies in Mexico, Argentina, Costa Rica, Peru” (p. 394). In addition, **CLEAR LAC 2015** implemented the first Evaluation Week in **Mexico**, which has now become a major promotional event for evaluation in Latin America and has been expanded to a global event in 2019 (p. 395). **CLEAR LAC** is therefore convinced that it itself “can reasonably be credited with a contribution to institutionalisation and capacity development for evaluation in the region” (p. 397).

Compared with the impact from the political system on the institutionalisation of evaluation, the contribution of civil society is very small and not very relevant. Even in the **USA**, evaluation is primarily perceived as a public task and civil society has not engaged in evaluation, particular not of its own accord. If evaluation plays a role in civil society, this had been clearly caused by the state and its demands for accountability and appropriate programme implementation. Civil society did not recognise evaluation as an instrument made for their own use—and there is almost no initiative to develop unique evaluation approaches for the third sector. There is no pressure on the political system to improve evaluation, evaluation use or evaluation quality—not even in countries like **Brazil**, **Canada**, **Mexico** and the **USA** where big and important civil society organisations with a fair amount of resources exist and a critical view on state activities is common. Causality is everywhere in the other direction: the increasing importance of evaluation in the political system spread over to civil society and included them by the demand for improving the implementation of public programmes and projects. While national authorities were held responsible by evaluations, they passed not only the programme management but also performance measurement and—nowadays increasingly—effect control into the hands of private actors.

This caused a diffusion of evaluation as management instrument always everywhere—but did not lead to ‘ownership’ for this tool in civil society.

Professionalisation is a process that seems to be widely decoupled from these dynamics in the political system. As mentioned before, the first impulse to implement study programmes in **USA** derived from public reforms and the need for evaluation skills—but this was before the real ‘boom time’ of evaluation. Surprisingly, the ‘take-off’ of evaluation in the 1980s/1990s in **USA** was not followed by a boom of new or advanced study programmes. Stagnation of numbers was accompanied by high fluctuation rates—almost no study programme established over time and not many new ones join the list in a sustainable manner. From the perspective of the system of professions, evaluation is still a niche, mainly seen as a specialisation of traditional subjects like economy, education, psychology, public policy or sociology. The transdisciplinary character of evaluation obviously hinders the establishment as a discipline by its own—even in the **USA** where evaluation raised at a time, where most of the social sciences were still young and growing. The institutionalisation of evaluation in the system of professions is not driven by the political sector and the increasing demand as one should expect. There are other barriers and drivers within the system that are of certain importance.

One of the key drivers in the system of professions seems to be a very individual one: professors and lecturers are free to develop individually specialised courses and study programmes. Such initiatives may consolidate over time but they are still linked to the person responsible. If this person leaves the university or retires, this offer may vanish if it is not installed in the study programme of the university and the university recognises this area as a special field of expertise. Especially small subjects may be replaced for political reasons, giving room for other subjects. Obviously, evaluation failed to take a particular hurdle: from being a highly specialised part of an existing subject to being an established programme supported by a couple of renown experts, giving the university some reputation within the academic world. There are probably some examples like the Master of Evaluation in **Costa Rica** or the Graduate Courses by the Evaluation Center at Western Michigan universities, but in general this hurdle stopped the further development of evaluation at **American** universities at an early state.

### *Challenges*

A number of reasons are cited in the case studies for a lack of use of evaluation, among other poor qualifications of evaluators and missing quality standards, resulting in non-appropriate evaluations, or—on the demand side—simply missing political interest. For example, the **Argentinian** report states, and that also applies for other countries, that evaluations are not in great demand and a vision is missing “that gives rise to a policy of capacity building, promotion and use of evaluation” (p. 48).

This is particularly true for civil societies which do not engage in evaluation and not pushing towards using this instrument or towards a higher quality of the evaluations executed by state organisations. To add for **Argentina**, civil society “evaluations usually focus on the performance of the programme or project, assessing results and drawing lessons that allow the improvement of the NGO’s or funding agency’s general programming. Sometimes, these evaluation practices are carried out by inner members of the organisation, based upon their knowledge management area” (p. 55). Evaluation as such is seen as a chore, demanded by government or donors and used by them for controlling the performance of project or programme implementation. If there is any use of evaluations within civil society organisations, then it is for management reasons and to learn for improving their own performance. This perception is widespread in many civil society organisations and not limited to Argentina.

However, even in countries where evaluation practice is strongly developed, such as in the **USA** and some **Latin American countries**, it can be observed that “a major barrier to use evaluation was stakeholder rejection of findings based on personal beliefs rather than data” (USA, p. 366f.). Evaluation activities are still regarded “as a resource drain on already limited funding and a distraction from their case mission” (USA, p. 366).

Although there is a well-established self-assurance in civil society to be an important actor for social development—and large, powerful civil society organisations and foundations are existing (particular in **USA**), there is no evidence for any unique tradition and methodological development of evaluation—in opposite, there is some evidence that non-profits are more averse to evaluation than federal agencies (cf. Carman & Fredricks, 2008; Carman, 2009).

Another obstacle is that in many Latin American countries evaluation is still understood as a “control mechanism” (**Bolivia**, p. 76; **Costa Rica**, p. 251; **Ecuador**, pp. 286f.) and thus has a negative image. This is most

explicitly brought down to the point in the **Costa Rican** case study by using the following words “that it tends to be thought that the nature of the evaluation refers to measurement of skills for a position, or the development of personal roles that could end in a sanction” (p. 251). This is particularly the case because evaluation is often used primarily for accountability purposes. The Costa Rican report states that one of the greatest challenges for evaluation is: “to change from a culture of accountability to one based on evaluation as a tool for improving public management” (p. 246). Despite the state of development and institutionalisation, evaluation is still perceived as a “bureaucratic requirement” (**Ecuador**, p. 285) and “there is a threat of over formalisation, which could result in a more bureaucratic environment” (**USA**, p. 363).

While looking at the Monitoring and Evaluation systems in Latin America, **CLEAR LAC** says that these “institutional frameworks and legal structures were anchored in a high-turnover bureaucracy, in the absence of a strong civil service tradition” (p. 397).

While controlling is seen as the main task of evaluation, civil society does not use this instrument for holding public authorities responsible for their work or the impact of public programmes or projects. Besides the widespread accountability focus that is also common in civil society organisations, “there is still a lack of knowledge and training on evaluation and clear methodologies that allow for their participation” (**Costa Rica**, p. 255). This may also lead to poor evaluation quality and evaluation is often seen as a “waste of time and resources without added value” (**Ecuador**, p. 285)—both in public authorities and in civil society.

This is a result of another challenge: the poor interlinkage between the system of professions including academic teaching and research on evaluation and the practical needs of the political and the social system. While some countries offer professional study programmes, this does not necessarily lead to more or better evaluations—and the other way around: complaints on evaluation quality from public authorities or civil society organisations do not result in pressure on universities or other training providers to offer better training programmes meeting their needs. As a result, most people doing evaluations “are not evaluation experts; rather they are experts in other fields that due certain circumstances have learned some evaluation concepts and methods” (**Ecuador**, p. 290). This is particularly true for countries with a poor or non-existing offer of professional training programmes. However, the development of such kind of

trainings is left to the academic sector without any trials for cooperation towards professionalisation.

Another difficulty for the systematic use of evaluation results is fragmentation, which occurs not only in countries where evaluations are conducted decentrally by individual ministries (e.g. **Argentina**, p. 50), but also where evaluations are centrally regulated by planning ministries or other institutions (e.g. **Costa Rica**, p. 247). Especially in countries where the practice of evidence-based policy is well advanced, evaluation faces a new challenge by competing with other instruments, for example performance management, auditing and data analysis (**USA**, p. 362). In this context, fragmentation means a lack of profile of evaluation and ambiguous usage. For most Latin American countries, the **CLEAR LAC** article in this volume stated: "that the M&E systems in Latin America are legally formalised for the most part, this does not amount to permanent anchoring and/or consolidation" (p. 400). **CLEAR LAC** identifies the following as central challenges: "deficits in human capital (which leads to low quality of evaluations), the underutilisation of evaluation for decision making, lack of citizen interest and legislators' involvement, as well as problems of fragmentation and lack of coordination of the entire management cycle" (p. 400). Their report also states that "the region as a whole advanced in the adoption of results-based management tools, including monitoring, evaluation, as well as performance evaluation-based budget, performance audits and governmental metrics of impact and development progress, with varying degrees of success and institutionalisation" (p. 396). This statement is a clear indication for the challenges already described here, namely that evaluation is perceived primarily as an instrument of control and is closely linked to budgeting processes, but that other functions of evaluation, such as learning or enlightenment, are missing.

These cumulative challenges extracted from the case studies form a vicious circle, including bureaucratic implementation and practice, limited knowledge and competences, and critical distancing to evaluation as a control mechanism. Each of these components is driving evaluation forward but in a rather limited and mostly unpopular way. The enlightenment and development potential of evaluation is merely overlooked.

Therefore, the future of evaluation is uncertain in an increasing climate of alternative facts that do not require empirical evidence, but claim validity solely on the basis of the vehemence and volume with which



they are presented. A discipline that creates evidence for decision-making, secures it with scientific thoroughness and makes it transparent and thus accessible to public discourse, can have a rather disruptive effect on the formation of political opinions. Alternative concepts—less challenging for ideologies and political beliefs—may overcome evaluation and its opportunities for continuous evidence-based learning.

This is particularly true for countries where populist governments are in power (such as the **USA, Mexico, Brazil, Bolivia, Ecuador**) and which cultivate a posttruth leadership style. But the existence of a “general anti-expert sentiment in public opinion” also represents a “great threat to the institutionalisation of evaluation” (CLEAR LAC, p. 401). This does not only concern countries where evaluation is weak anyway, but also countries in which evaluation is most strongly institutionalised and has been used to date. Above all, in the **USA**, where a president has ruled for the past four years, for whom lying was part of political practice and a legitimised instrument for pushing through particular interests, this also put pressure on evaluation. One of the most prominent evaluation experts in the USA, Michael Quinn Patton, recently proclaimed that culturally and politically evaluation science is under attack (Patton, 2018, p. 184).

“The emergence of anti-science trends in the **USA**, such as ‘alternative facts’, ‘fake news’ and a ‘posttruth world’, threaten evidence-based cultures that are sacrosanct to the evaluation community” (p. 359). Nevertheless, the authors of the case studies in this volume conclude that: “Evaluation practice in the USA is vast and robust. Evaluation is going on in just about every sector of our society” (p. 374). This may be caused by the strong and very active American Evaluation Association, providing both competences and networking—and at least indirectly some lobbying. Although “the public fails to recognise evaluation as a profession” due to a “lack of a professional designation” and credentials (p. 371), there are many activities below surface supporting the institutionalisation of evaluation done by AEA (e.g. the establishment of evaluation centres or coordinators at public agencies, p. 360).

In **Mexico**, the government that took office in 2018 disapproves of the performance evaluation system implemented by the previous government, and there have been fierce disputes between CONEVAL and the government. However, contrary to the announcement that CONEVAL’s budget would be reduced by 50% and personnel by 20%, only marginal cuts have been made so far. Nevertheless, the case study states: “this system continues because it is in the regulatory framework (...) there

have been no substantial changes in the Evaluation System". The authors conclude: "This shows institutional strength beyond changes in national policy" (p. 321).

In **Canada**, political influence on evaluation seems to leave a greater mark. A decade of Conservative Party ruling (2006–2015) has led to the fact that "the Canadian landscape of evidence-based policy-making was bleak" (p. 154). A serious change is then reported: in 2015, Chouinard and Milley (2015, p. 14) stated that "the actual use of evidence seems to be in decline, while the infrastructure surrounding evaluation (... the evaluation industry) seems to be on rise". Nowadays, some 'evidence-producing organisations' were disbanded.

Nevertheless, using stable, legally secured evaluation structures seems to ensure to survive political attacks and continue evaluation practice. At least, until the legal and organisational framework conditions are not fundamentally changed. Unfortunately, this is not enough for an enriching and fruitful development of evaluation towards a more appropriate and accepted use.

## REFERENCES

- Alvarez, S. E., Rubin, J. W., Thayer, M., Baiocchi, G., & Laó-Montes, A. (Eds.). (2017). *Beyond civil society: Activism, participation, and protest in Latin America*. Duke University Press.
- Busso, S., & De Luigi, N. (2019). Civil society actors and the welfare state: A historically-based analytical framework. *Partecipazione e Conflitto—The Open Journal of Sociopolitical Studies*, 12(2), 259–296.
- Carman, J. G. (2009). Nonprofits, funders, and evaluation: Accountability in action. *The American Review of Public Administration*, 39(4), 373–390.
- Carman, J. G., & Fredericks, K. A. (2008). Nonprofits and evaluation: Empirical evidence from the field. *New Directions for Evaluation*, 119, 51–71.
- Chouinard, J. A., & Milley, P. (2015). From new public management to new political governance: Implications for evaluation. *Canadian Journal of Programme Evaluation*, 30(1), 1–22.
- Derlien, H.-U. (1990). Genesis and structure of evaluation efforts in comparative perspective. In R. C. Rist (Ed.), *Programme evaluation and the management of government: Patterns and prospects across eight nations* (pp. 147–176). Transaction Publishers.
- Encarnación, O. G. (2001). Civil society and the consolidation of democracy in Spain. *Political Science Quarterly*, 116(1), 53–79.

- Fernandes, T. (2015). Rethinking pathways to democracy: Civil society in Portugal and Spain, 1960s–2000s. *Democratization*, 22(6), 1074–1104.
- LaVelle, J. M., & Donaldson, S. I. (2016). University-based evaluation training programmes in the United States 1980–2008: An empirical examination. *American Journal of Evaluation*, 31(9), originally published online 13 January 2010.
- Mark, K., & Pfeiffer, J. F. (2011). *Monitoring and evaluation in the United States Government: An overview* (ECD-Working Paper series No. 26), Washington: IEG.
- Meyer, W. (2016). Towards professionalization? The contribution of university-based training programmes in pioneer countries. In R. Stockmann; & W. Meyer (Eds.), *The future of evaluation: Global trends, new challenges, shared perspectives* (pp. 98–112). Palgrave Macmillan.
- Micheletti, M. (2017). *Civil society and state relations in Sweden* (2nd ed.). Routledge.
- Patton, M. Q. (2018). Evaluation science. *American Journal of Evaluation*, 39(2), 183–200.
- Rist, R. C., & Paliokas, K. L. (2002). The rise and fall (and rise again?) of the evaluation function in the U.S. Government. In J.-E. Furubo, R. C. Rist, & R. Sandahl (Eds.), *International atlas of evaluation* (pp. 225–247). Transaction Publishers.
- Segsworth, R. V. (2002). Evaluation in the twenty-first century: Two perspectives on the Canadian experience. In J.-E. Furubo; R. C. Rist; R. Sandahl (Eds.), *International atlas of evaluation*. Transaction Publishers, pp. 175–190.
- Stockmann, R., & Meyer, W. (2013). *Functions, methods and concepts in evaluation research*. Palgrave Macmillan.
- Wiarda, H. J. (2018). *Civil society: The American model and third world development* (2nd ed.). Routledge.
- Wollmann, H. (1984). Emergence and development of policy research in the U.S.A. In G. Thurn (Ed.), *Development and present state of public policy research: Country studies in comparative perspective*. Wissenschaftszentrum.
- World Bank (2020). *GNI per capita, atlas method (current US\$)*. <https://data.worldbank.org/indicator/NY.GNP.PCAP.CD>. Accessed 15 September 2020.



## The Institutionalisation of Evaluation in Europe and the Americas: A Comparison

*Reinhard Stockmann and Wolfgang Meyer*

To conclude, the attempt should be made to compare the results of the institutionalisation of evaluation in Europe (cf. Stockmann et al., 2020) and the countries of the Americas in order to bring out similarities and differences.

There are some methodological problems to be considered:

The obligation to use a uniform analysis grid has led to comparable results in both the **Europe** and the **Americas** study. However, it should be noted that, of course, this could not prevent the fact that not enough information and data are available for all *requested* items in the individual countries analysed. The interviews additionally accomplished by the authors concerning the questions for which there was only a meagre

---

R. Stockmann (✉) · W. Meyer  
Center for Evaluation (CEval), Saarbrücken, Germany  
e-mail: [r.stockmann@ceval.de](mailto:r.stockmann@ceval.de)

Department of Sociology, Saarland University, Saarbrücken, Germany

W. Meyer  
e-mail: [w.meyer@ceval.de](mailto:w.meyer@ceval.de); [w.meyer@mx.uni-saarland.de](mailto:w.meyer@mx.uni-saarland.de)

© The Author(s), under exclusive license to Springer Nature  
Switzerland AG 2022

R. Stockmann et al. (eds.), *The Institutionalisation of Evaluation in the Americas*,  
[https://doi.org/10.1007/978-3-030-81139-6\\_17](https://doi.org/10.1007/978-3-030-81139-6_17)

supply of information could only change very little in some cases. This is particularly true for activities within the social system, where the differentiation and diversity showing up here make it difficult to present a good overall picture.

Although all authors of the **Europe** and **Americas'** case studies were provided with a glossary of the terms used during the study in addition to the binding analysis grid, this did not always lead to a consistent use of the terms. This was generally not due to the authors, who were aware of these differences, but rather to the different ways in which these terms are anchored in the countries themselves.

In spite of these methodological challenges, which arose equally in **Europe** and the **Americas**, the results were hardly affected by this. This is mainly due to the fact that the effects described are so strong that even terminological imprecision cannot impair them.

If one considers the huge cultural, linguistic, social, economic and political differences within the continents and even more between them, it is astonishing that the comparative perspective of the three systems analysed reveals more similarities than differences.

However, a comparison of the legislative and organisational anchoring of evaluation in the political system reveals one serious difference. While in **Europe** evaluation is anchored in national laws in only three (of 16) countries under study (in three of these even in the constitution) and national decrees exist in only about half of them, the picture is significantly different in the **Americas**: only three countries (out of 11) have no national evaluation laws and all have national decrees.

There is also a difference with regard to the role of evaluation in the national audit offices: while in **Europe** in about half of all countries the audit offices explicitly also carry out evaluations, in the **Americas** there are only three, the **USA** and **Canada** and in all of Latin America only **Colombia**.

If the user side is being analysed, it is the other way around. Measured by the sectoral spread, thus the number of sectors or policy fields in which evaluations are regularly carried out, and measured by the intensity and frequency with which evaluations are performed, it becomes clear that **Latin America** shows a large implementation deficit. In other words, although there are far fewer legal and organisational evaluation structures in **Europe**, the degree of utilisation is significantly higher in Europe than in **Latin America**.

However, if one looks at the use of evaluation in national parliaments, the image again is quite uniform. To put it bluntly: evaluation results play a role in very few parliaments. In **Europe**, only **Switzerland** represents a certain exception, and in the **Americas** this is the **USA**.

If the four indicators for the legal and organisational anchoring of evaluation are combined to form the dimension 'institutionalisation' and it is correlated with the three indicators for the 'use of evaluation', then at best a moderate correlation can be seen, with striking differences between **Europe** and the **Americas**.

In **Europe**, there are four countries (out of 16) that have a very low (rated zero) utilisation rate: **Italy**, **Poland**, **Portugal** and **Romania**. These countries also have a very low degree of institutionalisation (between 0 and 0.25). In **Latin America**, there are even five countries (out of nine) that have an utilisation level in the political system that is rated zero: **Argentina**, **Bolivia**, **Brazil**, **Costa Rica** and **Ecuador**. But unlike in Europe, the degree of institutionalisation was rated much higher in these countries.

The positive benchmark in **Europe** is formed by **Switzerland** and **The Netherlands**, which both have a high degree of institutionalisation (0.88 each) and utilisation (1 respectively 0.83). In the **Americas**, the **USA** and **Canada** are outstanding (institutionalisation: 1 each, utilisation 1 respectively 0.83).

In **Europe**, **Finland** and **Germany**, with a relatively high degree of institutionalisation (0.75 and 0.5, respectively) coupled with a medium to high degree of utilisation (0.5 and 0.67, respectively), are still in the top group. Between these four best performers on the one hand and the bad performers (**Italy**, **Poland**, **Portugal** and **Romania**) on the other hand, there is a broad group of eight countries with average values.

In the **Americas**, this group is completely missing. Between the bad performers with a utilisation rate of zero and the best performers **USA** and **Canada**, there is a second small top group with relatively high institutionalisation values: **Chile**, **Colombia**, **Mexico** and **Peru** (all over 0.63) and, with the exception of Colombia and Peru, also relatively high utilisation rates (over 0.67: Mexico, Chile).

There are enormous differences in **Latin America**. This is primarily because there are a number of countries that have a high degree of institutionalisation but a very low degree of utilisation, such as **Costa Rica** (difference 0.63), **Ecuador** (difference 0.63) and **Colombia** (difference 0.55). In **Europe**, in no country, except **Latvia**, the difference is greater

than 0.33. In other words, even when looking at individual cases, this gap between the degree of institutionalisation and the degree of utilisation is striking.

Overall, this means that there is at most only a medium correlation between the degree of institutionalisation and utilisation, which is considerably less strong in **Latin America** ( $r = 0.60$ ) than in **Europe** ( $r = 0.82$ ).

In general, the following conclusions can therefore be drawn that the legislative and organisational anchoring supports the extent to which evaluation is used in the political system, but the institutionalisation of evaluation is far away to guarantee its use. Furthermore, the other way round, it can be stated that there is no country in which there is no legal framework and yet evaluations are carried out to a large extend.

As already mentioned above, the differences within the social systems are huge caused by varied traditions and heterogeneities within civil societies. Therefore, the communalities are surprising: neither the size, degree of organisation, financial capacities, political influence and engagement nor the philosophy, spirit or cultural diversity do influence the contribution of civil society for developing evaluation in the countries—it is everywhere almost close to zero. While there are some new and specific concepts and approaches that derived in the environment of civil society evaluations, these innovations are merely introduced by evaluators and not by civil society institutions. In fact, there are no institutions to anchor evaluation in civil society practice in opposite to state authorities and even the professional organisations are primarily driven by academics and private think tanks or state actors like administrations, implementing organisations, ministries or parliamentarians. Although some civil society organisations join them as members, one cannot identify them as a major driving force in developing such organisations.

Especially in the beginning, evaluation is seen as a threat or at least as an attempt by state authorities to control civil society and therefore some resistance occurred. In most cases, the state forced the adaptation of this instrument more or less strictly by setting up rules and regulations for civil society organisations. Step by step, evaluation convinced a growing but still very small number of civil society organisations as a management tool for organisational learning and has been established in particular fields of practice (especially in education, social services and development cooperation). They developed their own monitoring and evaluation institutions and practise within their own organisation, adapted to the organisational

culture and the needs for collective behaviour. Although civil society organisations are emphasising more on qualitative methods and participative approaches in comparison with state authorities, this did not flow into a separate institutionalisation of evaluation (e.g. focused networks or shared principles of use)—at least none of the authors in the **Americas** or in **Europe** reported this. There is no significant effect on institutionalisation of evaluation by civil society demand even in countries with a strong contribution of civil society for shaping and developing social life by providing services or bundling requests to the political systems.

Situation is a bit different if it comes to the supply side, the offer of evaluation services and the development of institutions for professionalisation of these services. Most of these activities are strongly linked to the academic system of the country and these systems developed their own rules for adapting new professions. From a global perspective, the system of professions in the **USA** has been dominating the formation of new academic disciplines and influenced the establishment in other countries, both in **Europe** and in **Latin America**. In most cases, global academic networks are dominated by **US** organisations and individuals with strong linkages to Western Europe and these networks are setting the agenda both for research and for teaching through scientific associations, journals and academic standards.

Evaluation is not an exception from this rule: the first academic institutions for evaluation derived in the **USA** and already in the 1970s, several universities offered courses on evaluation and implemented even whole new study programmes. With a certain delay (but still in the 1970s), a couple of evaluation journals were installed and the Joint Committee on Standards for Educational Evaluation (JCSEE) formed to develop systems of standards for evaluation practice in education. In these early years, evaluation became an important element of well-established disciplines like education or psychology and within the global academic network the idea of evaluation as part of the discipline spread over to **Western Europe**, particular **Germany, Scandinavia, Switzerland** and **UK**. Moreover, other disciplines like economy, public administration and policy, public health and sociology discovered the instrument and implemented it as a part of their own discipline. This increased the diffusion of evaluation at universities particular by forming study programmes, specialised think tanks and research networks with a strong focus on evaluation.

This diffusion process within the academic system seems to be almost independent from the development on the demand side. There are several



countries both in **Europe** and in **Latin America** which do have evaluation study programmes but almost no demand for evaluation from the political or social system. And one can also find the opposite: countries with a certain demand for evaluation but no academic infrastructure on evaluation research or training. The institutionalisation of evaluation in the academic system is remarkably poorly linked to the institutionalisation in the political and the social system—and therefore to the formation of an evaluation market and the forces of demand and supply.

One important element of professionalisation is the foundation of networks and associations to organise the communication between the members of a new forming profession. In most cases, the origin of such kind of organisations is also the academic system and the development of the largest evaluation organisation, the American Evaluation Association (AEA) in the **USA**, followed this pathway—although the increasing membership is caused by the successful inclusion of non-academics and internationals.

Hence, this is not the usual way in **Europe** and **Latin America**. In **Europe**, the evaluation organisations are much smaller but the most relevant ones—the European Evaluation Society (EES) and the **German**-speaking DeGEval (*Deutsche Gesellschaft für Evaluation*)—are also formed by academics—with some early attempts to include commissioners and practitioners from the political and social system. Compared to AEA and the Canadian Evaluation Society (CES), these organisations remained rather small and the organisational degree of evaluation is much smaller in **Europe** than in **North America**.

In **Latin America**, the evaluation networks are more informal and less organised. Although some of them—especially *Red de Seguimiento, Evaluación y Sistematización de Latinoamérica y el Caribe* (ReLAC), the Latin American regional network and the **Brazilian** one—achieved a much larger size than the **European** organisations, they are less powerful and it is difficult for them to organise collective action in a regular way. There is a significant influence from internationals by supporting the networking for evaluation in **Latin America**, and therefore, the networks are less dominated by academics.

As an applied science, evaluation has to link both academics and practitioners, as well as to cross borders between scientific disciplines to form a profession by its own. These are difficult tasks and even within the evaluation community, there are some doubts whether evaluation should become a profession or stay as part of other scientific disciplines.

Moreover, there is a certain tendency of practitioners to decouple evaluation practice from academic requirements, making evaluation faster and cheaper—in other words ‘quick and dirty’. Evaluation as a profession has to overcome these tensions and especially in **Latin America**—but also in many **European** countries—the organisational power and the willing of evaluation communities for lobbying are very low.

While the increasing demand for evaluations in most countries in the **Americas** and **Europe** keeps this kind of discussions alive, it does not necessarily influence the development of professionalisation processes at universities and professional organisations. There is still a lack of obligatory standards for improving the quality of evaluation and to increase utilisation both within organisations and within societies as a whole. Evaluators do not need to have special skills or knowledges because calls for proposals usually demand for experiences in evaluation and not for competences proofed by certificates (and this excludes young academics with new and innovative ideas from the market). Commissioners focus on the needs of their own organisation and are merely not interested in a general development of evaluation community. Evaluation network managers provide opportunities for exchange but do not understand themselves as advocates for an arising new discipline. Engagement in evaluation at universities as well as in evaluation networks depends on voluntary individual behaviour—there are no formalised and somehow institutionalised pathways towards professionalism.

The most advanced institutions can be found in **Canada** where both the Treasury Board of Canada and the Canadian Evaluation Society are supporting institutions for quality assurance and professional requirements for evaluation practice. These are still isolated applications with rather limited diffusion effects. The evaluation market remains an open market purely linked to emphasises for improving evaluation quality by professional training and research on evaluation. Variations between countries are remarkably low—although there are huge differences in academic establishment, degree of organisation and quality of evaluation supply.

## DRIVING FORCES

First of all, it can be said that the **institutionalisation** of evaluation has made enormous progress in **Europe** and the **Americas** over the past decades.

This development has not been constant, but was spread over several waves. The **USA**, where evaluation structures were already established in the 1950s and 1960s, is a country of origin for evaluation. In **North America**, **Canada** did not join until the late 1970s; in **Europe**, the **Nordic countries** and **Great Britain** as well as **Germany** formed the frontrunners group in the 1960s and 1970s. In **Latin America**, **Mexico** rapidly developed into a star performer in the 1990s, accompanied by **Colombia** and **Chile**. In **Europe**, the Eastern and Southern European countries are among the latecomers or those that have hardly been able to establish adequate evaluation structures to this day, such as **Portugal** and **Romania**.

In **Latin America**, this group includes both states with a relatively high GNI per capita, such as **Argentina** (US\$11,200), **Brazil** (US\$9,130) and **Costa Rica** (US\$11,700), and states like **Ecuador** with a relatively low GNI (US\$6,080) through to the poorest in **Latin America**, such as **Bolivia** (US\$3,530) (World Bank, 2020). In other words, the economic strength or weakness of a country does not appear to be an explanatory factor for whether or not a country has a high degree of institutionalisation and utilisation of evaluation.

It can also be observed that especially in some of the former frontrunning countries the level of development has slowed down to a certain degree, in some cases even a high level, but has lost its dynamic (**Denmark**, **Finland**, **Germany**, **Switzerland** and **UK**). This development can also be observed in the **Americas**, especially in the **USA**, **Canada** and **Mexico**.

The question that remains to be clarified is what has driven the institutionalisation of evaluation in **Europe** and **Americas** in recent decades?

Regarding the political system, the answer is surprisingly simple and clear. Despite the enormous political, economic, social and cultural differences within **Europe** and the **Americas**, and even more so between these continents, there is one clearly identifiable force, and these are the efforts to modernise and rationalise administration. The reforms involved had at least two roots: financial management and New Public Management or results-based management. The origins, which originate in financial management, are coupled with the intention of improving the state's financial control. Planning, Programming and Budgeting Systems (PPBS) were developed for this purpose. These systems were a variant of the then popular systematic management approaches which aimed to improve

system effectiveness and efficiency and fund allocation by linking explicitly defined organisational objectives to output and impact parameters.

The second source of reform resulted from the effort to improve performance and the process of service provision in the public administration. To achieve this goal, the concepts of New Public Management and output- and outcome-oriented management were developed.

The starting point for these reform efforts in the second half of the twentieth century was the prevailing view, supported by economists but also prevalent among the population, that the state was inefficient, too slow and not customer-oriented enough compared with the private sector. In order to improve the quality of government services, concepts from the private sector were imported into the administration.

Starting from the premise that the market is better placed than regulations to bring about efficient and effective service provision, competitive structures should also be created in the public sector and market instruments such as cost-performance analysis, performance measurement and benchmarking should be used. The idea of seeing citizens as customers, in other words as someone who uses a service provided by the administration, was downright revolutionary. In addition, this service has to meet certain quality requirements, which also require comprehensive quality awareness and management within the administration and throughout the entire public sector.

Most significant for the evaluation, however, was the turning away from the centuries-old control of the administration via input variables. This is realised by the assignment of different inputs (such as financial resources, personnel, equipment) in order to motivate the administration to become active in certain, determined fields of work.

*“As the usefulness of resource allocation is not necessarily examined on the basis of services rendered, results achieved, or even outcomes produced, efficient, goal- or even sustainability-oriented actions are hardly encouraged. On the contrary: traditionally in public finances, especially economical operation can even lead to resource cutbacks in the budget for the next financial year.*

*Another central element of NPM is therefore an orientation towards output and outcome. [...] This means that political management should align itself with output and outcome requirements. The underlying logic here assumes that the objective that state wishes to achieve is the outcome, rather than merely the operation of the administration”.* (Stockmann, 2008, p. 58)

However, the traditional instruments of administration were not sufficient for this type of control. For the development of the output, outcome- and impact-related data required now, new techniques and instruments were needed, such as those provided by evaluation. This was the breakthrough for evaluation. It is therefore no exaggeration to say: without this kind of administrative modernisation and the concept of making strategic political and administrative decisions based on a rational informational underpinning, evaluation would have remained insignificant.

This result applies both to the institutionalisation of evaluation in **Europe** and to the **Americas**. With one serious difference that still has consequences today: in the **Americas**, evaluation is much more focused on the rationalisation of the budgetary choices than in **Europe**. Evaluation was primarily assigned to the service of financial management and thus first and foremost to financial control, budget allocation and accounting. Thus, evaluation, like audit, is mainly used as a control instrument.

In **Europe**, the development of evaluation was not as closely linked to this tradition and necessities of financial and budget management, but more closely aligned with the concept of New Public Management and results-based management. Thus, other performance criteria such as goal attainment as well as output- and outcome-measurement were in the foreground.

Because of this control image, in many countries evaluation is still regarded as a threat rather than as an instrument for increasing the efficiency and effectiveness of political or administrative measures and rules—especially in **Latin America**, but also in some **European** countries.

Therefore, in many **Latin American** countries, evaluation is still used in financial control and budget allocation processes rather than for ‘improving’ the policies of programmes. At best, project and programme managers are using evaluation as a learning tool to intervene in implementation processes. And evaluation is least used for enlightenment processes that create transparency between politics and society and open a fact-based dialogue.

Civil society organisations are a certain corrective because they use evaluation more towards this direction. Hence, they are nowhere driving forces for the development of evaluation and its institutionalisation, although some countries (e.g. **Brazil, Mexico and USA, Denmark, Finland, France, Germany, Switzerland and UK**) are characterised by a strong and very engaged civil society. The social system, in general, is

not very important for the institutionalisation of evaluation and did not develop an own set of institutions for evaluation practice and use.

This result does not only apply to **Latin America**, but to a lesser extent also to **North America** and **Europe**. Even there, evaluation is still frequently considered a control that can be accompanied by negative sanctions—rarely with beneficial effects.

Another driving force for the development of evaluation in **Europe** is the European Union (EU). By making the performance of evaluation one of the mandatory conditions of receiving subsidies from the EU structural funds, evaluation comes to be practised in those countries which had not previously applied the concept. This is especially true for the Eastern European states, which did not join the EU until after the millennium.

In **Latin America**, there is no comparable actor that could have exerted such a formative influence on the development of evaluation as the EU. However, various international and national donor organisations can be assigned a supporting role in establishing evaluation in Latin America. On the one hand, this was done by carrying out evaluations in the projects and programmes that were financed by them. This allowed best practice examples to be demonstrated and the idea of evaluation to be disseminated. In addition, targeted evaluation capacity development measures were also taken, for example via UN organisations such as the International Fund for Agricultural Development (IFAD), the World Bank and via bilateral development cooperation with various countries (including Germany) or other forms of cooperation such as university cooperations.

In most countries, however, international cooperation was mediated by governmental organisations, so that it can be summarised that the development of the evaluation is clearly policy-driven, despite all the stated differences regarding the context conditions in **Europe** and the **Americas**.

The central motor and input for the institutionalisation of evaluation was the politically induced implementation of PPBS to control administration—especially in the **Americas**—and the adaptation of concepts related to New Public Management and outcome-oriented or results-based management. In **Europe**, the EU proved to be another driving force, which is particularly lacking in **Latin America**. The international and national donor organisations were at best able to set an example

with their evaluation practice and their—all in all very modest—evaluation capacity building measures, but they were also not in the least able to develop the significant role as the EU institutions in **Europe**.

## CHALLENGES

Finally, a few challenges facing evaluation in **Europe** and the **Americas** will be identified:

- Since evaluation activities are primarily limited to the political system, there is a risk that evaluation is increasingly interpreted as a kind of technology of the bureaucracy turning in a rigid routine involving largely standardised and streamlined processes. These trends are reported from countries with highly developed evaluation systems in **Europe** and the **Americas**, such as the **USA**, **Switzerland** and **Germany**, and from those where evaluation is only marginally institutionalised, such as **Ecuador** and **Costa Rica** as well as **Romania** or **Portugal**.
- If there is a market for evaluation to be carried out externally, it is also reported that the routinisation and bureaucratisation of evaluation are increasingly changing the supply side, as more and more consultancies are emerging that process these orders as standard. As a result, evaluations are now only conducted primarily by private service providers with specific expertise but without well-founded knowledge in evaluation (**Switzerland**). Especially in **Latin America**, it can be seen that in a large number of countries internal administrative capacities have been built up leading to highly standardised audit-like examinations. Both developments result in a kind of impoverishment of evaluation, creativity and research precision gets lost, evaluation is increasingly reduced to a standard procedure that hardly generates any new knowledge and continues to develop innovatively.
- As already explained, the potentially available evaluation functions: Insight, Development (Learning), Control and Legitimation (Stockmann & Meyer, 2013, p. 74) are not used to the full, but are reduced to the control aspect. Here, too, it is apparent across continents—even though with an emphasis on **Latin America**—that in some countries evaluation is perceived as an annoying duty or even as a threat that leads one to expect negative sanctions. Moreover,

there is a lack of opportunities for participation because these evaluations are merely recognised as a state activity for control reasons. The inclusion of civil society actors in evaluations is very rare exceptions although many national authorities follow the directive to increase transparency and closeness to citizens. There is still a high degree of mistrust in some Latin American societies and civil society recognises evaluation as an instrument for state control.

As a result, many people develop an aversion to evaluation, which does not make them believe in an improvement of their situation by revealing weaknesses that need to be eliminated in order to increase the efficiency, effectiveness and efficiency of measures, programmes or legal regulations, but rather increases their fear of losing their own position. This is all the truer when there exists no culture of failure management in these countries and their institutions and criticism is interpreted as a personal attack.

- In many countries—and the case studies in **Europe** and the **Americas** have impressively shown this—there is a lack of interest in evaluation, both among civil society and in politics. There is a limited understanding of the potential and possible applications offered by evaluation. This can be observed even in countries where evaluation is firmly anchored and used as a tool in the political system, more precisely in administration. In addition, many organisations still hold the long-held belief that evaluation is a waste of financial resources that are better used for the actual purpose of the organisation. This even applies to the ‘best performers’ in the **USA** and **Switzerland**. Often, evaluation results are also underestimated because decision-makers prefer to be guided by their feelings or personal beliefs, or trust other sources of evidence.
- As a result, there is no specific ‘evaluation culture’ in the social systems with a different way of understanding, implementing and using evaluation than in national authorities and administrations. Particular institutions cannot be found in the social systems and the differences between countries and continents are surprisingly small. There is almost no difference between the **USA** with its long tradition of civil society engagement, including the awareness that citizens have to do things by themselves and should not wait for the state, **Switzerland** with its specific decentralised form of citizen inclusion, **Argentina** or **Romania** with a long history of repression against citizens activities or **Scandinavian** States with a cooperative



welfare state culture, linking citizen engagement and state action very closely together. Although evaluation is a bit more understood from a participative perspective and in an enlightenment lens, there is no pressure from civil society for developing evaluation towards this direction.

- There is one exception within the social system and this is the independent establishment of evaluation in the system of professions. Remarkably, the development on the supply side is almost independent from the one on the demand side. While the demand for evaluation increased in many countries for years, the offers of evaluation training, communication about research on evaluation, organisation of evaluation exchange and the quality assurance via standards and regulations do not follow this development. One can find countries with well-developed institutions for evaluation in the political system but no adequate institutions in the system of professions (e.g. **Mexico**) and the other way around good established evaluation in the system of professions with almost no institution-alised use in the political or social system (e.g. **Argentina**). Hence, the anchorage in the system of professions is still very weak and there is a high degree of fluctuation. Especially in **Latin America**, it was not possible to establish comparable strong evaluation associations like in **Canada** and **USA**, although such attempts started very early. The dominating form of such institutions are open networks with a high number of (merely passive) members and therefore not much power for organising collective action. In opposite to this, **Europe** developed a couple of well-organised evaluation associations with a lot of activities, but a smaller number of members—especially in comparison with the huge **North American** associations. There are more than 4,000 people organised in 19 national evaluation organisations in **Europe** (nine person per 1 Mill. inhabitants) compared to more than double the amount in **North America** (8,800 in two organisations, 24 person per 1 Mill. inhabitants) and approximately 9,000 (17 person per 1 Mill. inhabitants) in nine organisations in **Latin America**. However, neither the number of members in voluntary organisations for professional evaluation nor the amount of study programmes, training offers and research journals is somehow related to the evaluation market development in the countries.
- Evaluation is also faced with the inherent challenge that it is not a ‘protected’ procedure. Since evaluation can, in principle, be carried

out by anyone who feels called upon or entrusted with it, the quality of evaluations varies considerably. This may be bitterly disappointing for customers who, in principle, consider evaluation a useful instrument and want to use it for their decision-making processes. This cannot be prevented because evaluators, unlike other professional groups such as auditors, do not need to be certified or have an academic (or non-academic) title. In addition, the development of evaluation is undermined by mixing evaluation with other tools. For example, it is reported from **Finland** that evaluation is becoming increasingly detached from its traditional model and transmuting into a scattered sphere of co-creation and peer learning. The role of the evaluator then transforms into the one of a critical friend or co-designers. A similar tendency can be watched in **Denmark**, where evaluation is to an even greater extent being mixed up with other concepts such as audit, accreditation, quality assurance and in some countries also with organisational consulting. Mixing the different concepts does not only lead to loss of the distinct profile of evaluation—something that is not desired in professional politics—but evaluation is reduced to an instrument that is no longer clearly identifiable. This instrument then no longer fulfils its actual task of collecting, analysing and assessing data using empirical methods, independently and in accordance with scientific standards.

- A massive challenge for evaluation worldwide is the growing belief in alternative facts that do not require empirical evidence, but claim validity exclusively on the basis of the vehemence and volume with which they are presented. This applies not only to those countries in which populist governments are in power and which cultivate a post-truth leadership style. As the COVID-19 crisis shows, even broad strata of society in properly governed countries are vulnerable to absurd conspiracy theories, however illogical they may be. When evidence no longer counts for much, disciplines such as evaluation, which produces empirically proven facts with scientific thoroughness, are facing hard times. On the contrary, the **USA** and **Mexico**, to give two examples, show that evaluation can succeed even in times of political pressure on fact producers—at least as long as there is a legal framework for it.

Looking at this list of challenges, one might conclude that evaluation is an endangered species. Now, that is far from being the case. In many

countries in **Europe** and the **Americas**, the subject is booming, evaluation is regulated by law and firmly anchored in ministries, authorities and state organisations, and is used for decision-making processes. On the other hand, in many countries, evaluation does not yet have a solid foundation that is strong enough to withstand any political storms and societal earthquakes. Therefore, it is important to be aware of the challenges so that counteractive measures can be taken.

In our opinion, the Voluntary Organisations for Professional Evaluation (VOPEs) play an absolutely important role in this process. In many cases, they do not yet perform it because they are too busy dealing with themselves. This is mainly due to their self-image, as they see themselves as professional associations, mainly discussing technical topics with each other, organise conferences and workshops. They rarely publish magazines or develop professional standards.

But they are usually no lobbyists. This seems to be too offending to many, not compatible with scientific morals and professional ethics. But it is wrong being that cautious. If the evaluators themselves do not even support the idea that evaluation should play a central role in politics and society, who else should do it? A new approach is needed here! Evaluation associations must also act as promoters of evaluation. We must not stand aside. We must actively involve ourselves in politics and society with our demands for more evaluation and fact-based decision-making processes. Otherwise, we must not complain that we are not noticed.

However, not only the VOPEs, but also the universities have the duty to contribute to political and social enlightenment in addition to their educational and research tasks. The problem here is that the discipline does not play a major role within the universities of most of the countries under review. Only a few locations succeeded in establishing evaluation as a research study programme in an institutionally stable manner. In most cases, the study programmes and research activities are based on the commitment of individual professors, and when those leave, the corresponding offers are often terminated. Larger institutions with several evaluation professors, who are thus more firmly anchored at the university, can hardly be found at present. There is an urgent need for action to stabilise and better anchor evaluation within the academic field.

The development in **Latin America**, in particular, also shows that a formal, institutionalised anchoring of evaluation does not necessarily lead to increased, high-quality implementation. There is a certain tendency to 'work off' evaluation as a bureaucratic instrument, and its potentials

for political practice is not sufficiently used. It applies especially to the aspect of learning from evaluations, which is, at best, limited to one's own organisations and administrative units, but is not used in the sense of further developing the whole society.

This is particularly challenging civil societies, which should demand this from the actors of the political system.

Precisely here is where the evaluation community should start and, in dialogue with civil societies, should clarify the usefulness of evaluation for social development. Increasing the social acceptance of evaluation and its use for mutual learning will become the central task for the further anchoring of evaluation on both sides of the Atlantic. The COVID-19 crisis currently highlights the urgent need: political measures are losing rapidly social acceptance without production and use of knowledge about its impacts, balancing carefully the positive effects by costs and side effects during implementation processes. While people are heavily affected by the immediate consequences, they want to be included in decision-making processes or at least feel their own interests being respected when harsh decisions for combating the crisis are made. A lack of rationality and fairness in 'top-down' management leads many citizens away from democracy and to follow populist leaders with crude and simple solutions. Evidence-driven policy based on functional evaluation institutions has an important role to play in preventing these dangers.

## REFERENCES

- Stockmann, R. (2008). *Evaluation and quality development: Principles of impact-based quality management*. Peter Lang.
- Stockmann, R., & Meyer, W. (2013). *Functions, concepts and methods in evaluation research*. Palgrave Macmillan UK.
- Stockmann, R., Meyer, W., & Taube, L. (2020). *The institutionalisation of evaluation in Europe*. Springer Nature Switzerland/Palgrave Macmillan.
- World Bank. (2020). *GNI per capita, atlas method (current US\$)*. <https://data.worldbank.org/indicator/NY.GNP.PCAP.CD>. Accessed on 14 December 2020.