



*Comparative Policy Evaluation*

# **POLICY EVALUATION IN THE ERA OF COVID-19**

Edited by Pearl Eliadis, Indran A. Naidoo,  
and Ray C. Rist



“The diversity and richness of the topics show immediately how important the book is. The findings ... to name a few, [such as] the deprioritization of evaluation during the crisis, the taken-for-grantedness of a number of underlying assumptions, describes [the] many efforts of evaluators.”

**Frans L Leeuw**, *Professor Law, Public Policy and Social Sciences, Maastricht University*

*“The strength of the book is the combination of the perspectives and the relevant insights they present. The book is undoubtedly timely, relevant, and important, not only for understanding what happened during the past years, but also in preparing for what will come.”*

**Per Øyvind Bastøe**, *Associate professor, VID Specialized University, Oslo, Norway*



**Taylor & Francis**

Taylor & Francis Group

<http://taylorandfrancis.com>

# Policy Evaluation in the Era of COVID-19

Did evaluation meet the challenges of the COVID-19 crisis? How were evaluation practices, architectures, and values affected? *Policy Evaluation in the Era of COVID-19* is the first to offer a broad canvas that explores government responses and ideas to tackle the challenges that evaluation practice faces in preparing for the next global crisis. Practitioners and established academic experts in the field of policy evaluation present a sophisticated synthesis of institutional, national, and disciplinary perspectives, with insights drawn from developments in Australia, Canada, and the UK, as well as the UN.

Contributors examine the impacts of evaluation on socioeconomic recovery planning, government innovations in pivoting internal operations to address the crisis, and the role of parliamentary and audit institutions during the pandemic. Chapters also examine the Sustainable Development Goals, and the inadequacy of human rights-based approaches in evaluation, while examining the imperative proposed by some authors that it is time that we take seriously the call for substantial transformation.

Written in a clear and accessible style, *Policy Evaluation in the Era of COVID-19* offers a much-needed insight on the role evaluation played during this unique and critical juncture in history.

**Pearl Eliadis** is Associate Professor (Professional) at the Max Bell School of Public Policy at McGill University in Montreal, Quebec. She also has a private law practice with a focus on human rights, national institutions, and democratic development and has worked with the United Nations (UN), the European Union (EU), and the Organisation for Economic Co-operation and Development (OECD) in several sub-Saharan and Asian countries on a range of national and global projects at the intersection of evaluation and the rule of law. She is a member of the Quebec Bar Association and the Ontario Law Society and held the position of president of the Human Rights section of the Quebec Bar for three terms until 2021. She also lectures at McGill's Faculty of Law and is a full member of McGill's Centre for Human Rights and Legal Pluralism.

**Indran A. Naidoo** is the Director of the Independent Office of Evaluation of the International Fund for Agricultural Development (IFAD) Rome, Italy. As former Director of the Independent Evaluation Office of the United Nations Development Programme (UNDP), New York, he transformed the evaluation function in terms of independence, output, and engagement, producing over 150 country and corporate evaluations. The National Evaluation Capacity series grew to become the largest and most influential platform for advancing evaluation at the government level. Previous roles included that of Deputy Director-General of M&E at the independent Public Service Commission and M&E Director at the Department of Land Affairs. He served as Vice-Chair of UN Evaluation Group (UNEG), instructor at the International Program for Development Evaluation Training (IPDET), and was a founding board member of the South African Monitoring and Evaluation Association (SAMEA). He holds a DPhil and post-graduate degree in Geography and Education and is the recipient of numerous awards and extensively published.

**Ray C. Rist** is one of the creators and co-directors of the original IPDET and former president of the International Development Evaluation Association (IDEAS). Retired from the World Bank, Dr. Rist continues to advise organizations and national governments throughout the world on how to design and build results-based M&E systems. His career includes 15 years in the U.S. Government with senior appointments in both the executive and legislative branches. He has held professorships at John Hopkins, Cornell, and George Washington Universities and has been a Fulbright Fellow at the Max Planck Institute in Berlin. He has authored and edited 35 books, written more than 140 articles, and lectured in more than 85 countries.

## **Comparative Policy Evaluation**

Edited by Ray C. Rist

The Comparative Policy Evaluation series is an interdisciplinary and internationally focused set of books that embodies within it a strong emphasis on comparative analyses of governance issues—drawing from all continents and many different nation states. The lens through which these policy initiatives are viewed and reviewed is that of evaluation. These evaluation assessments are done mainly from the perspectives of sociology, anthropology, economics, policy science, auditing, law, and human rights. The books also provide a strong longitudinal perspective on the evolution of the policy issues being analyzed.



**Taylor & Francis**

Taylor & Francis Group

<http://taylorandfrancis.com>

# **Policy Evaluation in the Era of COVID-19**

**Edited by  
Pearl Eliadis, Indran A. Naidoo,  
and Ray C. Rist**



First published 2023  
by Routledge  
605 Third Avenue, New York, NY 10158

and by Routledge  
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

*Routledge is an imprint of the Taylor & Francis Group, an informa business*

© 2023 selection and editorial matter, Pearl Eliadis, Indran A. Naidoo, and Ray C. Rist; individual chapters, the contributors

The right of Pearl Eliadis, Indran A. Naidoo, and Ray C. Rist to be identified as the authors of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

*Trademark notice:* Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

*Library of Congress Cataloging-in-Publication Data*

Names: Eliadis, F. Pearl, 1959- editor. | Naidoo, Indran A., editor. | Rist, Ray C., editor.

Title: Policy evaluation in the era of Covid-19 / Edited by Pearl Eliadis, Indran A. Naidoo, and Ray C. Rist.

Description: New York, NY : Routledge, 2023. | Series: Comparative policy evaluation | Includes bibliographical references and index. |

Identifiers: LCCN 2022049291 (print) | LCCN 2022049292 (ebook) | ISBN 9781032452968 (hardback) | ISBN 9781032452975 (paperback) | ISBN 9781003376316 (ebook)

Subjects: LCSH: Political planning--Evaluation--Cross-cultural studies. | COVID-19 Pandemic, 2020--Cross-cultural studies. | Sustainable development--Evaluation--Cross-cultural studies. | Human rights--Evaluation--Cross-cultural studies.

Classification: LCC JF1525.P6 P652 2023 (print) | LCC JF1525.P6 (ebook) | DDC 303.48/57072--dc23/eng/20230105

LC record available at <https://lcn.loc.gov/2022049291>

LC ebook record available at <https://lcn.loc.gov/2022049292>

ISBN: 978-1-032-45296-8 (hbk)

ISBN: 978-1-032-45297-5 (pbk)

ISBN: 978-1-003-37631-6 (ebk)

DOI: [10.4324/9781003376316](https://doi.org/10.4324/9781003376316)

Typeset in Bembo Std  
by KnowledgeWorks Global Ltd.

# Contents

<i>List of Contributors</i>	xi
<i>Acknowledgments</i>	xiii
<b>Introduction</b>	1
PEARL ELIADIS, INDRAN A. NAIDOO, AND RAY C. RIST	
<b>1 What Does the Pandemic Mean for Evaluation?</b>	24
JAN-ERIC FURUBO	
<b>2 COVID Crisis: Time to Recalibrate Evaluation</b>	43
MARIA BARRADOS, STEVE MONTAGUE, AND JIM BLAIN	
<b>3 The Unbearable Lightness of Rights: Evaluation and COVID-19 Responses</b>	61
PEARL ELIADIS	
<b>4 Knowledge Production in a Pandemic: Supporting Accountability at Pace in the United Kingdom and Canada</b>	83
JEREMY LONSDALE AND MARIA BARRADOS	
<b>5 The Impact of the COVID-19 Pandemic on the Effective Use of Evaluation in Supporting the Sustainable Development Goals</b>	101
ROBERT LAHEY AND DOROTHY LUCKS	
<b>6 Implications for Evaluation, What We Learn from the UN and Country COVID-19 Response Plans, and Reflecting on Future Scenarios</b>	124
INDRAN A. NAIDOO	

<b>7 Do Lockdowns Work?: Evidence from the UK</b>	140
RAY PAWSON	
<b>8 The Role of Evaluative Information in Parliamentary Oversight of the Australian Government's Responses to the Pandemic</b>	160
PETER WILKINS	
<b>9 Evaluation for Systems Transformations: Lessons from the Pandemic</b>	183
MICHAEL QUINN PATTON	
<b>Afterword</b>	206
RAY C. RIST	
<i>Index</i>	210

# List of Contributors

**Maria Barrados** is currently the Executive-in-Residence at the Sprott School of Business and serves as Adjunct Professor, Department of Sociology and Anthropology, both at Carleton University. She is the former president of the Public Service Commission of Canada.

**Jim Blain** graduated from the University of London (UK) in economics before joining the Canadian federal public service where he held various positions in program evaluation. He has retired from this position as Director General, Program Evaluation, Human Resources and Skills Development, Canada, and has continued to consult government on evaluation issues for a number of years.

**Pearl Eliadis** is Associate Professor (Professional) at the Max Bell School of Public Policy at McGill University. She is also a lawyer with an international law practice in human rights, national institutions, and democratic governance. She has worked for organizations such as the UN and OECD on a range of issues at the intersection of evaluation and the rule of law.

**Jan-Eric Furubo** has worked at different Swedish ministries, governmental agencies, and The National Audit Office. He has published works about the role of evaluation in democratic societies and was co-editor of the *International Atlas of Evaluation* (2002), *Evaluation – Seeking Truth or Power?* (2011), *Evaluation in Turbulent Times* (2013), and *The Evaluation Enterprise* (2020). He was the first president of the Swedish Evaluation Society and has had different roles within the European Evaluation Society.

**Robert Lahey**, President of REL Solutions Inc., was founding head of the Canadian Government's Centre of Excellence for Evaluation. He has advised globally (World Bank, UN agencies, UNEG, specific countries, international networks) on capacity building of national monitoring and evaluation systems, and evaluation use in supporting implementation of the SDGs.

**Jeremy Lonsdale** was a Director at the National Audit Office, UK until 2021, latterly being responsible for defense value for money studies. He worked on secondment as a Senior Research Leader at RAND Europe

2014–2016. His publications include as co-editor of “*Performance Auditing: Contributing to Accountability in Democratic Government*” (2011).

**Dorothy Lucks** is the Executive Director of Sustainable Development Facilitation (SDF) Global, a social enterprise based in Australia, that supports strategic and evidence-based sustainable development processes globally, including active evaluation practice in complex contexts. Dorothy is a Credentialed Evaluator and Fellow of the Australian Evaluation Society.

**Steve Montague**, partner of Performance Management Network Inc., is a credentialed evaluator, a Fellow of the Canadian Evaluation Society, and has four decades of experience as a practitioner in performance measurement and program evaluation as a management consultant, an adjunct professor, and as an evaluation manager in a major Canadian federal government department.

**Indran A. Naidoo** is the Director of the Independent Office of Evaluation of the International Fund for Agricultural Development (IFAD). He has served as Director of the Independent Evaluation Office of UNDP and Deputy Director-General of the Office of the Public Service Commission of South Africa. Indran is a globally recognized evaluation thought leader, speaker, manager, and author, who has spearheaded international reform initiatives.

**Michael Quinn Patton** is Founder and CEO of the Utilization-Focused Evaluation organization, US, and has written several books on evaluation. He sat on the faculty of the University of Minnesota for 18 years and is also a former President of the American Evaluation Association. He continues to be an active trainer and workshop presenter.

**Ray Pawson** is the Emeritus Professor of Social Research Methodology at the University of Leeds, UK. He is best known for pioneering the “realist” approach to program evaluation. His books include *Realistic Evaluation* (1997), *Evidence-based Policy* (2006), and *The Science of Evaluation* (2013). A new monograph, *How to Think Like a Realist*, is in preparation.

**Ray C. Rist** was one of the creators and co-directors of IPDET and also a Former President of the International Development Evaluation Association (IDEAS). Retired from the Independent Evaluation Group of the World Bank, Dr. Rist continues to advise organizations and national governments throughout the world on how to design and build results-based M&E systems.

**Peter Wilkins** is an Adjunct Professor at The John Curtin Institute of Public Policy (JCIPP) at Curtin University. He served as Western Australia’s Deputy Ombudsman. Prior to this, he served as Western Australia’s Assistant Auditor General Performance Review. He is a National Fellow and Western Australia’s Fellow of the Institute of Public Administration Australia.

# **Acknowledgments**

Many thanks to the Independent Office of Evaluation of the International Fund for Agricultural Development (IFAD) for their aid in editing this text.



**Taylor & Francis**

Taylor & Francis Group

<http://taylorandfrancis.com>

# 6 Implications for Evaluation, What We Learn from the UN and Country COVID-19 Response Plans, and Reflecting on Future Scenarios

*Indran A. Naidoo*

## **Introduction and background**

Historically, evaluation has evolved in a context of supply and demand, with the form of evaluation and the priority accorded to it based on the requirements given. In the development context, international funding and donor interests drove the profession, with an emphasis on accountability. The forms of evaluation that measured results to inform policy, funding decisions, or both, drove the development of the practice. Such an approach situated evaluation as a form of audit for the recipients, who often perceived the process as critical and punitive rather than constructive and beneficial. This reluctance to engage meaningfully accounts in part for the slow uptake of evaluation by most countries globally, as the profession was initially informed also by the Eurocentric evaluation literature that failed to explain its value accountability. Misperception and misinformation persist today in the way international bodies use evaluation, effectively serving as an additional assurance or fidelity function to governance bodies (Schwandt, 2019).

The Paris Declaration on Aid Effectiveness<sup>1</sup> was an important shift, as it emphasized self-determination and sovereignty, and began a process where evaluation became less donor-driven and more country-driven. Through the efforts of the United Nations Development Programme (UNDP), its convening power has been used to stress how evaluation is linked to concrete outputs and outcomes. As a result, the receptivity to evaluation has improved and the emphasis on accountability has decreased (Wilton Park Dialogue, 2018). This, in turn, has increased the interest of governments in the potential of evaluation, given that its framing had shifted away from funding conditionality. Geo-political shifts, together with a recognition that evaluation can benefit countries by helping them improve their development effectiveness, have also contributed to evaluation being embraced over time. The time when a discrete evaluation report from an independent or credible evaluation office should be viewed as a definitive form of uncontested judgement has passed.



COVID-19 has brought another major shift to evaluation, as Jan Eric Furubo argues in his chapter in this book. The evaluative conversations in this new construct will be informed by multiple information providers and actors and occur across different platforms and modalities (Rist & Stame, 2006). The exclusivity held by evaluators, irrespective of reputation or credibility, will change as their voices will be but one of many informing evaluation conversations. This chapter examines some of the changes that have occurred and that currently influence evaluation, for the perspective of its reconfiguration in a post-COVID era, its engagement with other research providers who have entered the evaluative space, and a reflection on what this means for evaluation functions and professionals in the future.

### ***Evaluation emphasis changing from accountability to supporting SDG attainment***

Over the past two decades, and especially in the last decade, evaluation developed in two fundamental ways. First, it has moved from being something which international players and donors insisted on, to being driven at a country-level and by civil society actors. Second, evaluation's value proposition has transitioned from being merely accountability-oriented to supporting policy formulation and promoting learning, especially towards the 2030 Agenda for Sustainable Development ("Agenda 2030"). All countries are now actively engaged in the global evaluation-development space and provide multiple models of leadership by introducing new philosophies and forms of evaluation. As a result, there is a greater sense of ownership of evaluation at the country-level.

The UN's convening power has undoubtedly advanced the role of evaluation in promoting progress towards the attainment of the Sustainable Development Goals (SDGs), which has emphasized national leadership and ownership of evaluation and has sought to improve the quality of national evaluation systems. The National Evaluation Capacity (NEC), organized by the Independent Evaluation Office (IEO) of the UNDP, found that the majority of the 160 participating governments had improved their use of evaluation to advance towards Agenda 2030 (Naidoo, 2020b, pp. ix–xi) cumulatively over the last decade.

The UN engages formally with governments through its convening power, with the NEC series being a central platform. In these events, it has engaged most evaluation networks around key topics such as the SDGs, evaluation criteria, and methodology, and has provided training to participants. Networks included the Evaluation Cooperation Group (ECG),<sup>2</sup> the United Nations Evaluation Group (UNEG),<sup>3</sup> and the EvalNet.<sup>4</sup> The meetings have introduced contemporary thinking on evaluation matters and have helped build capacity for government participants.

Government participants used the NEC platform to share their experiences in using evaluation, to report on and advance the attainment of the SDGs (Naidoo, 2020b, pp. ix–xi). The discussions indicated the achievement of maturity over time concerning earlier work and themes. These themes began from evaluation foundation-building and setting evaluation policies to more substantive discussions on the role of evaluation and how it was used in a more grounded manner to advance development agendas, and measure and report progress towards the SDGs (Naidoo, 2010, pp. 303–320). Evaluation moved towards a focus on providing practical solutions (Schwandt, 2005, pp. 95–105). The result was increased capacity to support the attainment of Agenda 2030 (van den Berg et al., 2017).

The progress of the NEC series over the past decade reflected these changes, moving evaluation in a direction that was action-oriented and pragmatic, whilst supporting future planning. First, it explored themes such as evaluation as a public good (2009, Morocco), followed by evaluation and public policy (2011, South Africa), and then advanced discussions on the implications for principles of evaluation (2013, Brazil). It progressed from a focus on evaluation as a tool to demonstrating how good evaluation helped improve people’s lives (2015, Thailand). The discussion then incorporated how UN efforts to promote development through the SDGs was a way of measuring progress (2017, Turkey). The last event took stock of what was achieved in terms of evaluating SDG-evaluation attainment (2019, Egypt).

### ***Baseline pre-COVID and changes in the COVID-19 era***

In global discussions, the NEC has served as a forum for highlighting fresh ideas. In 2019, for example, it convened an event on good practice standards (Naidoo, 2020a, pp. 63–69) and a case study from the IEO helped countries reflect on their own evaluation evolution (UNDP, 2020). The forum identified the following building blocks of good practice, although changes are likely to occur in the new era, due to the reprioritization that will be inevitable as we move forward:

- Evaluation policy, including governance and funding;
- independence, objectivity, and the SDGs;
- quality assurance of evaluation (UNDP, 2019);
- collaboration between evaluation and audit (Naidoo & Soares, 2020);
- addressing substantive needs and demands;
- evaluation scope; and
- communicating evaluation (Universität Bern, 2018).

The extent to which governments have embraced these aspects of good evaluation practice vary, reflecting the differences in evaluation systems across and within countries. Although there has been progress over time, key aspects such as the structural independence of the evaluation units

were not always realized. Other factors also prevent optimal performance for evaluation functions, such as the lack of clear policies, reporting lines, and budgets. The evaluation models employed also varied, with some governments outsourcing to draw in evaluation capacity, whilst others relied on their own capacities. The IEO case studies are important in that they illustrate a high degree of variability in evaluation capacity, which did not place evaluation in a strong position during the COVID-19 crisis to be able to support needs with maximum impact.

***Country-level evaluation mechanisms may be challenged due to COVID-19***

Over the last two decades, significant effort and resources have been invested in developing evaluation principles, policies, and practices at the national level. The growth of evaluation has resonated well with progressive ideals of advancing and supporting democracy, transparency, and accountability. There has been significant uptake of the practice within countries and their governments following recognition of the potential value of evaluation in decision-making and performance improvement. The evaluation sector has responded by supporting the development of evaluation architecture within and across countries, with multiple complementary global efforts linking evaluation to the attainment of normative and development goals. This has helped infuse an evaluation discourse into the planning processes of governments and raise awareness about the importance of being able to measure and respond to results, whether derived from political or administrative commitments.

Substantial progress has been made to embed evaluation across all sectors. In particular, it has helped to build the practice through dedicated occupational categories for evaluation-related activities in governments. There have been advancements towards professionalization accompanying an expansion of evaluation networks and associations. There is an extensive dedicated literature on the subject, illustrated by the number of books, journal articles, and diversity of experience demonstrated in the multiple mediums. Along with the political, civic, and administrative systems that advance the practice, there are greater efforts to systematically build and use evaluation capacity. The demand for accountability also comes from citizens who wish to see credible reports of results (Naidoo, 2004, pp. 8–11). Numerous evaluation networks and associations reflect the priorities of different evaluation constituencies, including consultants and evaluation professionals, commissioners, government users, academia, and civil society. All share the common ideal that evaluation seeks to make a difference by improving performance. As part of oversight, and together with audit, evaluation has been driven by criteria that aim to optimize the use of resources, promote efficiency and effectiveness, measure relevance and sustainability, and create value (Naidoo, 2020c, pp. 177–189).

***Changing realities may alter how evaluation is conducted***

Some features of the pre-pandemic evaluation architecture do not align well with the requirements for information now needed by countries. The shifts during the COVID-19 crisis occurred because traditional evaluations were seen as increasingly outdated along with the actors who understood the evaluations. The shifts come from department- or agency-specific approaches towards *holistic country approaches*, using multi-sectoral integrated approaches within sector-specific decision-making, based on multiple streams of information (compared to traditional evaluative information, which is generally discrete and based on singular reports). Whilst the discussions during the pandemic suggested joint evaluations and information-sharing between departments and agencies (UNEG, 2020), the linkages to other departments or agencies, or evaluating “as one,” remains largely non-existent. This more siloed approach has limited the relevance of reports in all-of-government or all-of-society approaches, both of which are key principles stated in the UN COVID-19 socio-economic response (United Nations, 2020) in its efforts to “Build Back Better” (United Nations, 2020). The pooling of development resources to support recovery efforts assumes that evaluation capacities and resources should be blended. In practice this has not happened, as evaluation functions continue to operate in a siloed fashion, serving the more focused needs of various agencies and their governing councils.

These examples illustrate the shortcomings of an evaluation architecture that, despite its evolution in recent years, continues to be linear and simplistic, and assumes a high degree of predictability and stability. It also assumes regular funding flows premised on predictable budgets (including taxes, remittances, and Official Development Assistance), predictable growth rates based on historic trends, and overall optimism. Today, this predictability is lost, and the operational environment of evaluators has altered, as evaluators are now competing with new actors at the country level. Institutions with strong academic and research capacities have gained considerable traction in providing oversight services.

These academic and research institutions possess strong multidisciplinary networks and can produce comprehensive work of an evaluative nature. They may potentially challenge smaller evaluation units that do not possess such capacities or networks. Evaluation curricula have become more developed and strong support has been provided to build the skills of people who train as evaluators. Research institutions can draw on and harness the latest technology to access large databases needed for appropriate assessments of the scale and magnitude of development questions at the country level. Major research institutions are also able to draw on real-time and disaggregated data to conduct scenario planning. Government users need such information that can be provided at low or no cost in a rapid manner, and that can be focused on real socio-economic development

challenges (Voccia, 2021). In addition, technology has the capacity to replace the need for physical interviews and other ground-truthing, which may replace a key element that evaluators had not only used to support their professional role and legitimacy but also for verification and deeper understanding. This further reduces the opportunity cost of using national-level evaluation capacities over established evaluation outfits.

### **Findings from a review of the SERPs as it relates to evaluation**

This section of the chapter highlights some findings from the UN review of the socio-economic response plans (SERPs) and its the implications for evaluation in the future. These new review and planning processes, installed by the UN and government compacts across 140 countries, claim to be collaborative and work horizontally, and emphasize issues that should promote recovery like human rights and inequality. The UN review used a rubric to assess comprehensiveness and the extent to which the plans were data informed. It also examined economic performance and impact on population groups and focused on humanitarian crises, the environment, economic dependencies, and the impact of value chain disruptions.

#### ***What the plans seek to achieve and their focus***

The SERPs are joint government-UN documents, agreed by both parties. They seek to be comprehensive and emphasize joint responsibility for results. The policy and guidance documents intend to provide an empirical and logical basis for designing new development pathways. They seek to instil global normative values and priorities into the national sphere. The instrument claims to focus on response and planning efforts, and to be people-centred, whilst allowing countries to work out implementation modalities.

Each of the five pillars of intervention contains baseline data from which interventions may be monitored, and theoretical scenarios based on the severity of the crisis. The data to support the interventions is drawn from existing and planned studies to ensure that interventions are effective.

Discrete evaluation reports have little value in a collaborative context. Furthermore, evaluation entities do not have the ability to work within and address the comprehensive nature of the UN COVID-19 SERPs. This envisages a degree of joint leadership and funding for securing data, as well as developing a common understanding of what potential changes the crisis will require. The joint approach must also include an understanding and response to the humanitarian-environmental nexus and track deviations from SDG targets which could derail progress. Developing policy options to address vulnerabilities and inequalities with the intention to address structural inequality is also important.

The existing country-level evaluation approaches, even from the international evaluation offices, cannot provide the comprehensive approach required to deliver meaningful policy options. Evaluation has been largely absent in this development at the country level.

***There are indications that the traditional evaluation ecosystem shall change***

Whilst evaluation gained prominence for its potential role in supporting the attainment of the SDGs and progress towards meeting Agenda 2030, the COVID-19 crisis introduced new realities. Aspects of evaluation that were necessary for supporting the SDGs, such as a strong evaluation architecture, clear deliverables of products to inform SDG progress, and the resources to deliver these activities, no longer fit into a development planning paradigm. The new development priorities and more comprehensive ways of working triggered by the COVID-19 crisis rendered evaluation, in its current form, less effective. This is because evaluation is not configured to be agile and responsive and has generally worked by supporting discrete mandates or features that do not help in this new context. Nonetheless, strengthening local capacity for measuring progress on the SDGs remains important given the interlinked relationship between the SDGs and national development goals in many countries (UNDP, 2020).

***Government responses as gleaned from the SERPs***

The choice of shutting down the economy to save lives has been a source of tension and was hard to justify in the absence of economic measures to support the loss of incomes. Governments' ability to manage these conflicting goals was regularly challenged. It has brought to the surface questions of how well governments are able to address the humanitarian crisis. Whilst most claim that their response was evidence-based, the absence of sufficiently transparent monitoring and reporting systems means that health-protection plans are largely aspirational. There has been little publicly accessible evidence of progress based on what the plans have set out. The SERPs mention oversight committees, with collaboration and joint responsibility for results and reporting. Many government plans have included expanded membership to encompass academia, non-governmental organizations, and the private sector. Most SERPs include, as a minimum, the UN and government leadership at the country level, working to execute the plans jointly.

The extent to which governments respond to these factors, as described in the SERPs, will only be known through an independent monitoring and evaluation system with both national and international credibility to answer questions about the effectiveness of the COVID-19 measures. Examining whether scarce resources are targeted to the vulnerable and the

poor, whether new forms of economic activity have succeeded in reducing reliance, and whether “Building Back Better” does in fact happen, cannot be known based on inward-looking monitoring and evaluation systems. These are difficult political questions, especially as systems are unlikely to develop to the standard required given the political sensitivities associated with reporting transparently on government effectiveness during COVID-19.

The SERPs reports show that governments are already using the research capacity of universities to assist in the planning process, as they have access to other forms of data. However, one of the major challenges is the lack of data on key sectors of the economy, especially those most impacted by the pandemic, such as the informal sector, which makes up 60 per cent of the global workforce (an average figure, which is likely significantly higher in many lower-income countries) (United Nations, 2020, p. 17). The lack of disaggregated data by gender and other vulnerability markers has made it difficult to identify equitable solutions. These groups tend to be outside formal structures and face other levels of danger and vulnerability, such as discrimination and marginalization.

It cannot be assumed that information technology and internet connectivity will solve these data problems. There remains the problem of electricity, equipment, and connectivity costs; the digital divide is a hindrance in most countries. The COVID-19 crisis may have made the digital divide and many of its digital requirements worse. Working remotely is not an option for the informal and service sector and a nuanced approach to evaluation using virtual techniques would be required to include the most marginalized.

### ***What the SERPs suggest about an emerging evaluation architecture***

The SERPs indicate that the current oversight architecture is inadequate, as it has traditionally operated in a predictable rather than dynamic, crisis context. There is limited capacity for working across oversight structures or understanding that oversight can be a comprehensive process which is collaborative rather than mandate driven. These are multi-year national development plans, most of which contain references to monitoring and evaluation as a means to periodically assess progress. Whilst the plans are national in nature, they reflect a siloed approach of individual ministries, many of which do not collaborate.

As for UN interventions, they are evaluated by UN agency-specific offices, and results do not feed into a broader evaluation discussion. There has been limited UN agency collaboration and few efforts to change this through a new coordination system; sustained results are yet to be demonstrated.

The SERPs suggest that the shift taking place has an emphasis on more actors reporting on progress. These national actors acquire a crucial role

in country-driven evaluations and the evaluation practice suggested by the SERPs requires them to understand the scale of the changes promoted by interventions and to respond accordingly. The plans also emphasize the complexity of the current reality and the importance of taking it into account for policy development. Traditional evaluation practice, with its linear orientation and mandate-specific focus, is not agile enough to produce the insights necessary for the new context.

The existing evaluation architecture may no longer be relevant to shifting priorities and may be uncoordinated with new ways of working. If the SERPs are indicative of the future, there will be less emphasis on information from agencies and a greater focus on “conversation-driven,” collaborative and engaging work. The evaluation architecture will be less definitive and more focused on the future compared to the past.

All country development plans have been reframed to ensure their relevance to recovery efforts. The international community, too, has had to reassess how it measures its intervention success. The previous plans were purely focused on a development pathway for Agenda 2030 and the attainment of the SDGs (Naidoo & Soares, 2017, pp. 51–63). Now, the SDGs remain important but take on a new emphasis; joint guidance by the OECD and UNDP, for example, has positioned the pandemic as an opportunity to “spark a new wave of innovation and ambition” relating to Agenda 2030 and the SDGs (Independent Evaluation Office/United Nations Development Programme & Organisation for Economic Co-operation and Development/Development Assistance Committee, 2020).

### **The response by evaluators during the crisis with evaluation**

During the pandemic, pockets of evaluative activity have focused on alternative methodologies to support old practices rather than on recognizing the magnitude of the crisis and its future implications. The oversight contexts have changed alongside new demand and supply sources for evaluation. As previously noted, the academic and research sector has stepped into the space left by the inertia of evaluation leadership; with the limited visibility of evaluators on the frontline of an historic global crisis have, it is unclear what the future holds for monitoring and evaluation (The Wits School of Governance, 2020).

The review of over 80 SERPs (Naidoo, 2022) shows that the traditional oversight architecture no longer functioned as usual during the pandemic, but there has been significant growth in the offering from research institutions to governments to manage the response. There has been little contribution of evaluation expertise from the international evaluation networks to support recovery efforts, either through their agencies or collectively.



The fracturing of the stable environment that was conducive to an effective evaluation architecture and the inability of the international evaluation community to either respond creatively or recreate itself means this space is effectively lost.

***Evaluators should be able to adapt to the new development and evaluation discourses***

In the future, evaluators will need to expand their frame of reference and to understand and be able to work with complexity. Evaluators have tended to be reductionist in their approaches, simplifying complex issues in a manner that is hard to justify, using outdated methodologies, and generally being unable to evaluate beyond their agency mandates.

Evaluative skills require conceptualization at the global, regional, and country levels. They also require an understanding of the scale and interaction amongst various levels, being able to frame assessments in the context of political and developmental issues, and the ability to construct policy options. More specifically, and based on the SERPs review, evaluators must be able to frame the COVID-19 crisis globally against the backdrop of previous development trajectories and inherited vulnerabilities.

Each of the content areas includes a set of interconnections, which are complex and part of a fast-changing dynamic which is inherently political and influenced by geo-political factors. In addition to the factors already mentioned, there is the digital divide, the role of the diaspora in the context of population movements and migration, and changes to the operation of financial development institutions, including what the COVID-19 crisis means for debt and other obligations. Issues of food security, triggered by the closing of markets and disruption in production contribute to the complexity. Projecting ahead, major additional research capacity and streams of information will be needed, the most obvious being shifts away from singular agency or departmental evaluative reports for discrete audiences towards reports from established institutions. These should address the complexity and the nature of the comprehensive information required. This development emerges relatively well from think tanks and research institutions, which are also strong in providing multi-disciplinary perspectives.

***Evaluators need to understand and work with scale***

At the broadest level, the COVID-19 crisis amplifies existing inequalities and levels of differentiation. A response that is generic and presents an aggregated reaction will mask these inequalities and disproportionate impacts. To address this, however, would require sophisticated data and analysis, something not generally present in individual evaluation units. The emphasis on singular interventions, which is a feature of

agency-specific evaluation units (at the government level and those within the international evaluation networks) translates to their inability to deal with differentiated impacts across scale.

The SERPS review shows that interventions need to reach beyond urban communities to focus on any disproportionate impacts between and within peri-urban and rural populations. Many of the targeted populations do not benefit from public-sector infrastructure, making it difficult for them to access services. A key factor identified in the review is the digital divide, with the lack of electricity, funds, computers, and networks preventing remote education; tele-medicine; and other digital service provision. Therefore, the ability to evaluate across different geographic levels and scales means that the deeper levels of socio-economic differentiation are glossed over, in part due to the use of averages. Census data is not comprehensive enough to allow for pro-poor targeting. This means that most of the policy options and pro-poor policies will lack the benefit of solid insights. The reviews also point to increased discrimination based on gender and other grounds. However, in the absence of solid monitoring data, the real impact is unknown.

The implicit capital or headquarter bias in government operations is mirrored by urban and official data bias. Data tends to be aggregated and all the SERPs demonstrate a deficit in disaggregated data, even at the level of national census data. If information fails to highlight variations within the population, the policy responses will fail to address social and economic differentiation adequately. The policy options presented as interim responses in the SERPs have already shown a bias towards the aggregated data. Framing the response according to scale means moving beyond the comfort zones of the capitals and governments and generating information on historically marginalized. This is often not possible.

### ***Critical content areas that require specialized knowledge for monitoring and evaluation***

The COVID-19 crisis has multiplied the number of development challenges. Until the crisis, the SDGs served as a comprehensive set of common indicators to measure progress. The magnitude of the crisis resulted in much deeper changes that fundamentally affected established systems, including evaluation. It has been observed that the crisis paused regular activities, such as the established practice of reporting government progress against set plans, especially in countries with some form of democracy. The de-prioritization of this form of accountability, given the crisis context, has created questions which need to be asked if there is to be a reestablishment and reorientation of evaluation practice as an accountability measure.

*The reorientation of evaluation to work across sectors and agencies, producing high-quality real-time evaluative information for immediate recovery*

The evaluation sector was developed in environments which had a degree of predictability; the planning processes of governments assisted in fostering that stability, with a clearly established sets of users for results reporting. The demand and supply for this type of evaluative work, however, has taken place in silos, as mentioned above, and there has been very little horizontal collaboration in oversight. During the crisis, the established systems were interrupted. Resources were pooled and reprioritized, and reporting on results was no longer the sole preserve of any agency; rather, it became a joint collaborative reporting effort.

The SERPs mention evaluation, but with limited details, and many of the reviews come from non-traditional evaluation sectors such as research and academic think-tanks. They have been able to deliver at the speed and scale required. Good examples include the National Council for the Evaluation of Social Development Policy in Mexico and the National Institution for Transforming India (NITI) Aayog in India. They marshal national level evaluation capacities from academia as well as the public sector. Such institutions are also best placed to provide institutional legitimacy if required when it comes to making evaluative judgements. Whether the evaluation units of government or international agencies can contribute to this new space shall become evident over time.

Evaluation units tended to be small compared to these other entities, and the absence of significant responses during the pandemic likely indicates a lack of preparedness or an inability to retool evaluation to meet new needs and demands. Evaluation is not as familiar with big data and geo-spatial analysis tools and this gap has been evident in the reports produced by the academic and research sector. This work may not meet the standards of evaluation in terms of independence, but it has met research standards and has been able to deliver results.

The COVID-19 crisis has thus uncovered many weaknesses in the evaluation system. The focus on discrete interventions has no value in an all-of-government or all-of-society approach. This requires major ideological and behavioural changes from evaluation. Historically, evaluations have remained disengaged from policy and operational interventions; the new context requires more engagement.

*Evaluators need to work at multiple levels and be able to unpack disaggregated data*

There are three principal and interrelated challenges in implementing the response to COVID-19: equity, public sector capacity and data availability. In terms of *equity*, as the SDGs clearly articulate, it is of utmost

importance to reach all groups of people, especially those who are most vulnerable. In a post-COVID world, it becomes even more crucial given that the crisis has hit the informal sectors and vulnerable population the hardest. Implementing equity-focused programming requires robust *public sector capacity*. The review of the SERPs has highlighted weaknesses within the public sector infrastructure, which tends to be urban-biased. Thus, weak public sector capacity has an adverse impact on equity. COVID-19 has revealed the limitations of centrally driven, top-down approaches to programming and evaluation. The final factor, which is important for programming and evaluation in a post-pandemic world, is the availability of *reliable data*. This factor, linked to research capacity within countries, openness to alternative data sources and views, and media freedom is critical for evaluation.

*What is required of evaluators in the new context?*

At the country level, there has been at least some collaboration amongst the various agencies. A new UN system to improve coordination has shown to be effective, as reflected in the SERPs, which emphasize the pooling of resources, at the very least from the UN, towards a joint UN approach (Freeman et al., 2022). As the substance of the SERPs shows, the following attributes are required for any oversight and support function:

- Agility and the ability to work across mandates collaboratively and evaluate as one. The various pre-crisis efforts to bring about evaluation coordination to mirror the reform efforts seeking to get the agencies to work seamlessly were unsuccessful.
- Participation in the efforts for actual collaboration, reprioritization, and commitment to joint budgeting by agencies and departments, evidenced by the SERPs.
- Possession of specialized content knowledge and understanding required in the light of the new development context, before moving into developing monitoring and evaluation systems. There has not yet been an audit of the skills of evaluators against the new content focus areas at the country level. It is evident that the institutional capacity to provide the content proficiency and work to scale is more present in the academic and research fields than in the evaluation sector.
- Capacity for co-creation of knowledge and working in a holistic, all-of-society approach, which is generally not within evaluators' experience.

It was assumed that engagement would compromise the ability to provide objective judgement. The other shift in understanding is that past trends no longer offer any reasonable basis on which to offer propositions,

whether recommendations or insights. The magnitude of the crisis has been such that the focus is on immediate recovery efforts, working in challenging and under-resourced contexts with little time available to await long reports.

## Conclusion

The COVID-19 crisis has posed challenges and opportunities for evaluation. This chapter argues that the main response to these challenges has been superficial and methodological, employing stopgap measures to mitigate an inability during the pandemic to gather real-time, credible information that assists in decision-making. The use of tools like geographic information systems and remote sensing is part of a technological addendum to evaluation but cannot replace the need for what remains a strategic and analytical function (García & Kotturi, 2019). There have been shifts in the governance environment which has affected the evaluation architecture, which has been relatively secure for supporting fidelity evaluation. The need for classic accountability evaluation shall change as funding alters alongside geo-political shifts that call for more self-determination of evaluation. The inability of the evaluation community to adapt its value proposition and enter the new development space however is concerning and may affect its further relevance.

## Notes

- 1 The Paris Declaration was endorsed at the Second High Level Forum on Aid Effectiveness in 2005. It is a practical, action-oriented roadmap to improve the quality of aid and its impact on development (Organisation for Economic Co-operation and Development, 2005).
- 2 The Evaluation Cooperation Group is the professional network of the World Bank and regional banks.
- 3 The United Nations Development Group (UNEG) is the professional networks of the evaluation and oversight offices of the United Nations.
- 4 The DAC Network on Development Evaluation (EvalNet) is the evaluation network of the bilateral agencies and is led by the OECD Development Assistance Committee (DAC).

## References

- Freeman, T., Esser, A. L., & Vela, P. (2022). *Interim report: System-wide evaluation of the UNDS response to COVID-19*. United Nations. [https://unsdg.un.org/sites/default/files/2022-03/Interim%20Report-%20SWE%20of%20the%20UNDS%20Response%20to%20COVID-19\\_Final.pdf](https://unsdg.un.org/sites/default/files/2022-03/Interim%20Report-%20SWE%20of%20the%20UNDS%20Response%20to%20COVID-19_Final.pdf)
- García, O. A., & Kotturi, P. (Eds.) (2019). *Information and communication technologies for development evaluation*. Routledge.
- Independent Evaluation Office/United Nations Development Programme & Organisation for Economic Co-operation and Development/Development

- Assistance Committee. (2020, April). *Good practices during COVID-19*. United Nations Development Programme. [http://web.undp.org/evaluation/guideline/documents/covid19/IEOOECD\\_DAC\\_Joint-Guidance\\_COVID19.pdf](http://web.undp.org/evaluation/guideline/documents/covid19/IEOOECD_DAC_Joint-Guidance_COVID19.pdf)
- Naidoo, I. (2004, November/December). *The emergence and importance of monitoring and evaluation in the public service*. Public Service Commission News. [http://www.psc.gov.za/documents/pubs/newsletter/2004/psc\\_news\\_book\\_nov\\_2004.pdf](http://www.psc.gov.za/documents/pubs/newsletter/2004/psc_news_book_nov_2004.pdf)
- Naidoo, I. (2010). Monitoring and evaluation in South Africa: Many purposes, multiple streams. In M. Segone (Ed.). *From policies to results: Developing capacities for country monitoring and evaluation systems* (pp. 303–322). UNICEF. [http://www.theoryofchange.org/wp-content/uploads/toco\\_library/pdf/2010\\_-\\_Segone\\_-\\_From\\_Policy\\_To\\_Results-UNICEF.pdf](http://www.theoryofchange.org/wp-content/uploads/toco_library/pdf/2010_-_Segone_-_From_Policy_To_Results-UNICEF.pdf)
- Naidoo, I. (2020a, October). Architecture for evaluation effectiveness. In Independent Evaluation Office of UNDP. *Proceedings from the 2019 national evaluation capacities conference*. United Nations Development Programme. [http://web.undp.org/evaluation/documents/NEC/2019/NEC2019\\_proceedings.pdf](http://web.undp.org/evaluation/documents/NEC/2019/NEC2019_proceedings.pdf)
- Naidoo, I. (2020b, October). Foreword. In Independent Evaluation Office of UNDP. *Proceedings from the 2019 national evaluation capacities conference*. United Nations Development Programme. [http://web.undp.org/evaluation/documents/NEC/2019/NEC2019\\_proceedings.pdf](http://web.undp.org/evaluation/documents/NEC/2019/NEC2019_proceedings.pdf)
- Naidoo, I. (2020c). Audit and evaluation: Working collaboratively to support accountability. *Evaluation*, 26(2), 177–189. <https://doi.org/10.1177/1356389019889079>
- Naidoo, I. (2022). *Evaluation for transformational change*. Unpublished manuscript.
- Naidoo, I., & Soares, A. (2017). Incorporating the sustainable development goals in national evaluation capacity development. In R.D. van den Berg, I. Naidoo, & S. D. Tamondong (Eds.). *Evaluation for agenda 2030: Providing evidence on progress and sustainability* (pp. 51–64). International Development Evaluation Association. [https://ideas-global.org/wp-content/uploads/2017/12/IDEAS-web-REV\\_08Dec.pdf](https://ideas-global.org/wp-content/uploads/2017/12/IDEAS-web-REV_08Dec.pdf)
- Naidoo, I., & Soares, A. (2020). Lessons learned from the assessment of UNDP's institutional effectiveness jointly conducted by the independent evaluation office and the office of audit and investigation of UNDP. In M. Barrados & J. Lonsdale (Eds.), *Crossover of audit and evaluation practices: Challenges and opportunities* (pp. 182–193). Routledge.
- Organisation for Economic Co-operation and Development. (2005, March 2). *Paris declaration on aid effectiveness*. <https://doi.org/10.1787/9789264098084-en>
- Rist, R. C., & Stame, N. (2006). *From studies to streams: Managing evaluative systems*. Transaction.
- Schwandt, T. (2005). The centrality of practice to evaluation. *American Journal of Evaluation*, 26(1), 95–105. <https://doi.org/10.1177/1098214004273184>
- Schwandt, T. (2019). Post-normal evaluation? *Evaluation*, 25(3), 317–329. <https://doi.org/10.1177/1356389019855501>
- The Wits School of Governance. (2020, September 8). *The future of monitoring and evaluation* [Video]. YouTube. <https://www.youtube.com/watch?v=qTXhHbE3Fp4>
- United Nations. (2020, April). *A UN framework for the immediate socio-economic response to COVID-19*. United Nations Sustainable Development Group. <https://unsdg.un.org/sites/default/files/2020-04/UN-Framework-for-the-immediate-socio-economic-response-to-COVID-19.pdf>
- United Nations Development Programme Independent Evaluation Office. (2020). *Leaving no one behind: Evaluation for 2030. Proceedings from the 2019 National Evaluation Capacities Conference*. United Nations Development Programme Independent

- Evaluation Office. [https://nec.undp.org/sites/default/files/2021-03/NEC2019\\_proceedings-1.pdf](https://nec.undp.org/sites/default/files/2021-03/NEC2019_proceedings-1.pdf)
- United Nations Evaluation Group. (2020). *UNEG ethical guidelines for evaluations*. United Nations Evaluation Group. <http://www.unevaluation.org/document/detail/2866>
- Universität Bern. (2018, October 17). *IPDET 2018: Brown bag 2* [Video]. YouTube. <https://www.youtube.com/watch?v=71yV80jJqr4>
- van den Berg, R. D., Naidoo, I., & Tamondong, S. D. (Eds.). (2017). *Evaluation for agenda 2030: Providing evidence on progress and sustainability*. International Development Evaluation Association. [https://ideas-global.org/wp-content/uploads//2017/12/IDEAS-web-REV\\_08Dec.pdf](https://ideas-global.org/wp-content/uploads//2017/12/IDEAS-web-REV_08Dec.pdf)
- Voccia, A. (2021). Behind the glossy cover: Six reflections on communicating evaluations. *Independent Magazine* (1). [https://issuu.com/ifad\\_ioe/docs/independent\\_magazine\\_-\\_issue\\_1\\_web?fr=sZmE2MDM3MzgwMjU](https://issuu.com/ifad_ioe/docs/independent_magazine_-_issue_1_web?fr=sZmE2MDM3MzgwMjU)
- Wilton Park Dialogue. (2018). *Revisiting independence, objectivity, and the critically reflective role of evaluation for the SDG era*. <https://www.wiltonpark.org.uk/wp-content/uploads/2020/09/WP1600-Report.pdf>