MONITORING SYSTEMS IN AFRICA

Caitlin Blaser Mapitsa & Catherine Churchill

Monitoring Systems in Africa

Caitlin Mapitsa and Catherine Churchill

SUN PRESS

Monitoring Systems in Africa

Contents

About the Contributors

Caitlin Blaser Mapitsa is the research director and associate professor in monitoring and evaluation at the University of the Witwatersrand School of Governance. She has over 15 years of experience in monitoring and evaluation (M&E) research, organisational development, and capacity building. She is particularly excited about understanding what drives change in complex social systems and how to measure those changes. Her background has focused on interdisciplinary rural governance, and her work has been with organisations advocating in the inequality and climate justice sectors.

Rekgotsofetse Chikane is a public policy lecturer at the Wits School of Governance, political commentator, activist, and author of *Breaking a Rainbow, Building a Nation: The Politics behind the #MustFall Movements.* He is a graduate of the University of Oxford (Blavatnik School of Government), having completed his Master's in Public Policy degree in 2017, and he completed a Bachelor in Social Science (Hons) degree at the University of Cape Town in 2015, specialising in Public Policy, Public Administration, and Economics.

Caroline Nuga Deliwe is a doctoral candidate at Stellenbosch University and Chief Director responsible for Sector Planning, Monitoring, Evaluation, and Research at the National Department of Basic Education in South Africa. Caroline has over two decades of experience in education planning, management information systems, monitoring, evaluation, and strategy within the country. Caroline leads the strategy development, planning, monitoring, and evaluation of sizeable public sector programmes in response to long-range education development goals. Her work on management information piqued her interest in how educational institutions use data and information on learning outcomes to account for performance and improve development outcomes. Her research investigates how institutional arrangements in the schooling system influence accountability for performance and learning. She examines the constraints and benefits of the audit culture in government departments tasked with improving learning outcomes and how school accountability can strengthen information use at the school level to boost learning outcomes.

Catherine Churchill is an editor of postgraduate theses, academic tutor, course developer and facilitator. She is a development practitioner passionate about positive change through practical wisdom. Catherine graduated from the University of the Witwatersrand and completed various courses involving project management, policy development, capacity development, and results-based monitoring and evaluation. She has gained over 25 years of experience working in various senior management positions in the academic, public, and development sectors. She is piloting an international learning exchange programme between the University of

Tampere, Finland, and the University of Johannesburg, South Africa, focusing on Innovation for Sustainability, the Circular Economy, and "people-driven, place-based" local development.

Philip Browne has over 25 years of expertise consulting in the development sector and undertaking programme management, evaluations, policy analysis, and strategic planning with governments, development agencies, and NGOs in South Africa and the East, West, and Southern African regions. Philip is a gender and social inclusion specialist with extensive experience in gender mainstreaming for policy formulation, programme design, and developing gender-responsive M&E systems. He has an academic background in education (MPhil) as well as public and international development management (MPDM), along with Public Health (MPH), with a focus on social policy and inclusion, human rights-based approaches, social sector research, and political economy. Philip has extensive experience with conducting evaluations, social and health impact assessments, and gender, inclusion, and vulnerability assessments. He has specialised in applying specific evaluation and assessment tools, including gender analysis, gender lens investing, political economy, development investment prioritisation, and cost-benefit analyses. Philip has worked extensively in the areas of gender equality/women's empowerment, public health (particularly HIV and sexual and reproductive health), education, social development, economics and labour markets, and governance. He has experience in policy analysis, gender-responsive results-based planning and management, and developing theories of change (ToCs). Philip brings extensive policy, research work, and training in development assistance, human rights, gender, and sustainable development. He has strong analytical skills in project formulation and administration, programme implementation, and Monitoring & Evaluation. He has the capacity to produce detailed gender-responsive evaluations, research, policy, and programme documents. Philip is also an experienced trainer and facilitator in M&E, gender, social inclusion, and human rights. He uses a human-centred design methodology in all capacities of development interventions, and his clients have included government departments, UN agencies, the Global Fund, the World Bank, academia, international NGOs, global philanthropies, and the private sector.

Professor David Everatt is a former Head of School at the University of the Witwatersrand School of Governance, and is Principal Investigator for GRT-INSPIRED, an HDSS in Gauteng.

Abdoulaye Gounou is a senior M&E professional. He has been Director General of evaluation at the Prime Minister's office in Benin. He is currently the Head of the Bureau of Public Policy Evaluation and Government Action Analysis at the Presidency, Republic of Benin. He holds a diploma of Advanced Professional Studies from Senghor University of Alexandria, Egypt, in business administration and project management, and a Master of Law from the National University of the Republic

of Benin. He is involved in conducting process and impact evaluation of public policies, capacity development events, and collaborative projects development and implementation such as Twende Mbele, a Programme with South Africa and Uganda and WACIE Programme, the West Africa Capacity building and Impact Evaluation, a Programme with 3ie within the WAEMU countries in West Africa. The programme received financial support from WAEMU Commission and BOAD. Previously, he acted as Technical Adviser to Ministers (Prime Minister, State Minister of Planning, and Minister of Finance) from the Government of Benin for more than seven years in economic strategy, reforms, and private sector development.

Halfdan Lynge is a Senior Lecturer at the University of the Witwatersrand School of Governance. His research explores the effects of politics on public policy, i.e. how politicians' cost-benefit calculations affect ordinary citizens' lives. In addition, he is interested in data science and its application to public policy. Halfdan holds a doctoral degree from the University of Oxford, UK, and a Master's and Bachelor's degree from the University of Copenhagen, Denmark. He founded Sauti, a public opinion research company specialising in mobile-based surveys, and is a Civic Tech Innovation Network (CTIN) reference group member. Halfdan previously worked for five years with the UN in Africa and Asia. His last position was as Head of the UN Resident Coordinator's Office in Dhaka, Bangladesh, where he supervised a team of advisors to the UN Country Team. Halfdan started his career as a governance specialist and continues to work as a governance consultant for the UN, the World Bank, and non-governmental organisations.

Steven Masvaure is a Senior M&E Technical Specialist at the University of the Witwatersrand's Centre for Learning on Evaluation and Results. His research interests are governance, public participation, climate change, and Global South socio-economic development issues.

Pundy Pillay is a Professor of Economics and Public Finance at the University of the Witwatersrand School of Governance. He earned a PhD in Economics and an MA in Economics from the University of Cape Town. He completed his BA Honours at the University of the Witwatersrand. Before joining the school, he was a Visiting Professor at the Universities of the Western Cape and Johannesburg from 2007 to 2010. Between 2001 and 2007, he was a Senior Economist at RTI International. He was Head of the Policy Unit in the Office of the President between 1998 and 2000 and the Financial and Fiscal Commission Executive Director between 1995 and 1997. He began his academic career at the University of Cape Town as a Senior Lecturer in Economics from 1983 to 1994. Recent research of his in the local government sector includes '*Developing a Pro-poor Bus-Fare Policy for the City of Johannesburg*', '*A review of Community-Based Planning*' (also for the City of Johannesburg), and '*Transport Planning and Planners for the Local Government SETA*'.

Nedson Pophiwa is a senior lecturer in monitoring and evaluation at the University of Witwatersrand's School of Governance. During the last decade, he has been involved in several research and evaluation projects within the South African region. He holds a PhD in Economic History from the University of KwaZulu-Natal, an MA in Forced Migration Studies from the University of the Witwatersrand, an MA in African Economic History, and a BA Hons in Economic History, both from the University of Zimbabwe. Regarding his contribution to research, Nedson has published book chapters, peer-reviewed publications in journals, and has read papers at academic conferences. He is a co-editor of the *Evaluation Landscape in Africa: Context, Methods, and Capacity*, published by African Sun Media (2019), and *Africa in a Changing Global Environment: Perspectives of Climate Change Adaptation and Mitigation Strategies in Africa*, published by the African Institute of South Africa (2013).

Jamie Robertsen is a monitoring, evaluation, and learning expert with 12 years of experience in the development sector, specifically in monitoring and evaluation. Jamie is a PhD candidate and sessional lecturer at the University of the Witwatersrand School of Governance and has an MSc in Development Evaluation and Management from the University of Antwerp, Belgium. Jamie has worked across a broad range of sectors, including climate change and renewable energy, youth employment, governance, private sector development, social justice, infrastructure, and financial services. He has worked with varied stakeholders, ranging from non-governmental organisations to governments and donors, multilateral organisations, and philanthropic foundations. Jamie has worked in Bosnia, Herzegovina, Botswana, China, Ethiopia, Kenya, Namibia, Madagascar, South Africa, Zambia, and Zimbabwe.

Khotso Tsotsotso is currently a Director for MERL at Data Innovators, as well as a PhD candidate and sessional lecturer at the University of the Witwatersrand's School of Governance. Among his many interesting projects, he has led the development of the Evaluation of Ethical Guidelines for the Department of Planning Monitoring and Evaluation (DPME) in South Africa. He has also led the development of a research strategy and policy for the Transport Education Training Authority (TETA). He wrote TETA's Sector Skills Plan for four consecutive years from 2016 onward. Khotso led the design and development of the M&E plan for the Botswana Government's new TVET policy in partnership with the German Development Corporation. One of his many notable career highlights has included being recognised by the South African Monitoring & Evaluation Association (SAMEA) for his outstanding contribution to the practice of M&E as a young and emerging evaluator in Africa. He has both partici-pated in and led over twenty programme evaluations.

Acknowledgements

Acronyms[1]

APNODE	African Parliamentarians' Network on Development Evaluation
B-P	Budget-Program
CAA	Autonomous Amortisation Fund
CAD	Development Assistance Committee
CAD	Departmental Administrative Conference
CaR-GBAR	Results-based Budget Management Reform Framework
CASPPP	Analysis and Synthesis Unit of Project Performance and Programmes
CDMT	Medium-Term Expenditure Framework
CGP	Programme Management Committee
CIEPP	Institutional Framework for Public Policy Evaluation
CMR	Results Measurement Framework
CPRB	Budget Reform Steering Committee
CREAM	Clear, Relevant, Economic, Adequate, Monitorable
CS	Supreme Court
CSE	Monitoring-Evaluation Unit
CSPEF	Economic and Financial Programs Monitoring Unit
DD	Directorates Sectoral Development
DPDM	Department of Planning and Municipal Development
DPP	Directorate of Planning and Prospective
DG-ODD	Directorate General of SDGs Coordination
DGE-OCS	Directorate General of Evaluation and Observatory of Social Change
DGIFD	Directorate General of Investments and Financing of Development
DGPD	Directorate-General for Development Policies
DGSPP	Directorate-General for Monitoring Projects and Programs
DGTCP	General Directorate of the Treasury and Public Accounts
ECOWAS	Economic Community of West African States
GAR	Results-based Management
GIZ	German International Cooperation
3ie	International Initiative for Impact Evaluation
LOLF	Organic Law related to Financial Act
MDAEP	Ministry of Development, Economic Analysis and Prospective

---

1 Acronyms utilised in Francophone contexts are defined in English.

M&E	Monitoring and Evaluation
MEF	Ministry of Economy and Finance
NES	National Evaluation System
OECD	Organisation for Economic Cooperation and Development
PDC	Municipality Development Plan
PCC	Credit Consumption Plan
PC2D	Growth and Sustainable Development Programme
PAG	Government Action Plan
PAI	Annual Investment Plan
PIP	Public Investment Program
PPM	Procurement Plan
PTA	Annual Work Plan
RBM	Result-Based Management
SIAPIP	Integrated System of Analysis and Programming of Public Investments
SCRP	Growth for Poverty Reduction Strategy
SIGFiP	Integrated System of Public Finance Management
SHISEPIP	Harmonised and Integrated System for Monitoring and Evaluation of Public Investment Projects and Programmes
SMART	Specific, Measurable, Attainable, Relevant and Time-bound
SPAT	Planning and Territorial Development Department
SPICED	Subjective, Participatory, Interpreted, Cross-checked, Empowering, or Disaggregated
STDL	Planning and Local Development Department
SDGs	Sustainable Development Goals
TEF	Financial Execution Rate
TEP	Physical Execution Rate
TIB	Investment Dashboard
TM	Twende Mbele
UEMOA	West Africa Economic and Monetary Union (WAEMU)
UNDP	United Nations Development Program
UGPs	Project Management Unit
VNRs	Voluntary National Reviews
WACIE	West Africa Capacity and Impact Evaluation

Introduction

Caitlin Blaser Mapitsa

African governments invest significant resources in monitoring and compliance-related activities, hoping to improve the effectiveness of development interventions. However, this effectiveness has not always been realised, and many chapters in this book will explore the barriers between effective monitoring practice and effective governance and development interventions. As an illustrative example, a recent survey of officials within the City of Johannesburg found that over 70% of officials' time was spent on monitoring activities. With monitoring consuming such a significant amount of public sector resourcing, it is critical to understand better the purpose, structure, and function of monitoring systems in Africa.

In the M&E field, monitoring is often seen as the lesser companion to evaluation – a technical process of churning out data that can then be used for higher and more important strategic processes of planning and judging worth. However, while evaluation practice and institution building within African public sectors and research about its efficacy have enjoyed a decade of rapid growth, there is simultaneously growing admission that while evaluation is important, monitoring has roles in addition to feeding the evaluation processes. It can be used as an interceding step to planning and adaptive management without a formal evaluative step taking place. Furthermore, it also contributes directly to critical reporting, performance management, budgeting, and accountability processes. From this perspective, looking at monitoring systems as a crucial piece of organisations themselves becomes imperative.

This volume presents a holistic approach to monitoring systems. It presents a range of purposes that monitoring practices serve, followed by a selection of the components of a monitoring system, and then presents case studies of monitoring systems. The aim is to explore monitoring as a practice on its own, related but not constricted to its framing in performance management or evaluation. The hope is that through drawing together a range of practice-based cases that bring different perspectives to the practice of monitoring, academics, policymakers, and those

responsible for operating different pieces of a monitoring system will be able to better locate their piece of the monitoring puzzle with a more holistic view of the role monitoring should play in governance and development.

In **Section 1**, the first four chapters broadly explore three different purposes for which monitoring data is often used. The three chapters consider the strategic reasons for organisations investing in monitoring, the potential of monitoring to lead to development outcomes, and the interceding organisational structures and processes that can but do not necessarily lead to the effective use of monitoring. These chapters also consider specific trade-offs that need to be made in linking the design of monitoring systems to their intended purpose.

The chapters in **Section 2** consider four different components of monitoring systems. The first looks at how monitoring systems are embedded in organisational processes and considers different theoretical lenses for understanding how monitoring can contribute to organisational change. The next chapter focuses on indicators; while no discussion of monitoring systems is complete without a detailed discussion of this controversial part of monitoring, this chapter departs from the frequent practice of dictating a specific, technically correct approach to using indicators, and rather problematises the different choices that are made in the process of developing and interpreting indicators. The third chapter looks at how monitoring links to external 'big' data sources. It considers the implications of bypassing the process of interpreting and judging data entirely and creating a direct link between data and decisions. Finally, in the fourth chapter of this section, a South African case study concerning budgeting, accountability, and outcomes is explored through the lens of the monitoring practices seeking to advance best-practice in performance and programme-based budgeting systems.

In **Section 3**, three chapters provide case studies of monitoring in various contexts. The first examines multi-institutional partnerships and the promises and challenges of matching data to multiple organisational needs and systems of use. It also focuses on spatial data, which is unique in its contribution to planning. The second considers a national monitoring system's significance for a whole-of-government monitoring practice. The final chapter looks at frontline service delivery monitoring (FSDM) and the contributions citizen participation can make to public sector systems.

Section 1

THE PURPOSES OF MONITORING SYSTEMS

Chapter 1: Monitoring Systems in a Context of Complexity and Uncertainty

Caitlin Blaser Mapitsa

In times of uncertainty or rapid change, governments must have useful and practical monitoring systems underpinning their policy development. The COVID-19 pandemic has highlighted how rapidly dramatic social change can happen and how monitoring practice linked to effective research systems can anchor policy responses to the best available data, saving lives and mitigating adverse public health and economic and other societal impacts. Climate change's natural and social devastation is another realm where rapid scientific advances can critically shape policy responses (Yung, Louder, Gallagher, Jones, and Wyborn 2019; Vigoda-Gadot and Mizrahi 2016). Public sector policy-making processes exist in an increasingly uncertain context, and rigorous monitoring can provide information about these responses at a scope and speed unthinkable until recently (Nalubega and Uwizeyimana 2019).

Monitoring is the systematic process of collecting information about the implementation or results of a programme. It is often done through the use of indicators to determine programme performance. Monitoring can focus on any part of the programme, from implementing activities to achieving results or spending against the planned budget. Monitoring has many purposes, including contributing to evaluation as part of an M&E system, accountability and compliance, or performance improvement.

Monitoring is usually carried out internally as a core function of programme management, although occasionally, there is an element of external validation oversight. Monitoring usually focuses on data generated by the programme, whether it is organisational data, stakeholder-related data, budget data, or similar. Monitoring usually happens continuously and is linked to nested reporting levels. Usually, specific indicators or other types of programmatic data are gathered on a monthly basis, quarterly, and annually. Things monitored frequently, such as every month, are activity-oriented and speak to low levels of a results chain. In contrast, things monitored annually link to higher levels of a programme's results chain, such as a broader outcome. These results take longer to achieve but are also rarely exclusively reliant on one programmatic intervention, but need to triangulate with other stakeholder interventions, changes in context, and confounding variables.

Monitoring practices are primarily descriptive or observational and record and check the completed activities. In a training programme, this could include activities such as gathering and collating attendance registers from workshops to determine when they were held and how well-attended they were. A school feeding scheme could include checking the number of children fed. Monitoring a service delivery point, such as a clinic, could include information such as the demographics of people who used the service, the average wait times to receive the service, and the levels of satisfaction with the service.

Since monitoring could, but does not necessarily, include elements of judgement or analysis, programme monitoring activities do not usually require a broader context of programme logic or a theory of change that would explain why certain things are done or what they intend to achieve. However, as other chapters in this book will illustrate, the quality of monitoring may be enhanced with this context.

This chapter will explore how monitoring practice connects to organisational evaluation and learning functions and how monitoring effectively feeds into evidence use through evaluation systems. It will also look at the constraints to generating monitoring data that meets the needs of evaluative practice and organisational systems for adaptive management.

Monitoring Within and Beyond M&E

Over the last decade, the field of monitoring for development has shifted beyond focusing on monitoring as key to evaluation and performance management to taking a more systemic view of monitoring. This shift in perspective has defined monitoring in the context of development and reform and has further located monitoring within a range of cognate management tools, practices, and systems. These include accountability, performance budgeting, and evaluation, including data systems, reporting, and performance management systems (Wholey 2007; Van der Waldt 2004; Mayne and Zapico-Goni 2017).

Monitoring serves various purposes, with reporting, compliance, and accountability among the most important. Additionally, generating a foundation of data for evaluation is also a significant objective of monitoring. Much of the contemporary research on monitoring and evaluation assumes that these two practices are intertwined and referred to as "M&E" (Kusek and Rist 2009). While there has been a recent surge in literature focusing on evaluation systems and data management in African contexts, monitoring systems have not received the same level of attention. This lack of attention may stem from the literature's emphasis on the relative neglect of evaluation, particularly in the public sector, despite the substantial investments and resources allocated to establishing and improving monitoring systems. Consequently, research on evaluation has served, in part, to advocate for increased

evaluative practices and greater utilisation of the monitoring data generated (Porter and Goldman 2013). While this is undeniably important, it is equally crucial and inadequately addressed regarding a knowledge base development that can enhance the current significant investment of organisational time, attention, and resources in monitoring (Wholey 2007).

The literature on M&E systems in African public sectors broadly agrees that governments place overwhelming emphasis on monitoring practice (rather than evaluation) for public sector management in the region (Basheka and Byamugisha 2015; Mapitsa and Korth 2017). While the reasons and drivers for this (over)investment and the ends to which it contributes are contested, the phenomenon itself is relatively undisputed. Given its prominence, it is surprising that more resources do not exist linking empirical examples from public sector practice to current theory on monitoring systems. As part of this discussion, we underscore the importance of a systemic approach to monitoring, defining the varied purposes of monitoring, and looking at some of the enablers and constraints of monitoring practice. We will highlight the theoretical and practical implications a systemic approach to monitoring has on practice and problematise different components of monitoring systems from a range of disciplinary perspectives. The result aims to speak to both development and public sector audiences across the continent who use and craft the monitoring tools needed for effective governance.

A Systems Approach to Monitoring

The starting point for considering monitoring is often reporting on progress towards key performance indicators or a similar practice that may be described with different vocabulary based on organisational norms. As a result, a considerable amount of literature about monitoring focuses on the indicators themselves – how to develop them and how to measure them in a collectively understood way (Booth and Lucas 2002). While this is certainly an important part of monitoring (and is discussed at length in Chapter 5), it is a partial and incomplete view of the practice.

To examine how progress is reported, it is crucial to consider that monitoring takes place within a context of numerous interrelated functions, which serve as components of monitoring systems. The figure below illustrates six of these components. Some of these components, such as data systems, provide the inputs for monitoring practices. Others, such as evaluation practices, are situated at the use end of monitoring. However, all these factors influence how monitoring is conducted. They determine the incentives for monitoring and drive the underlying value systems that shape what will be monitored, how monitoring will be conducted, and by whom (Mihaiu 2014). Understanding these factors is key to grasping the essence of monitoring, as well as the facilitators and constraints of effective monitoring.
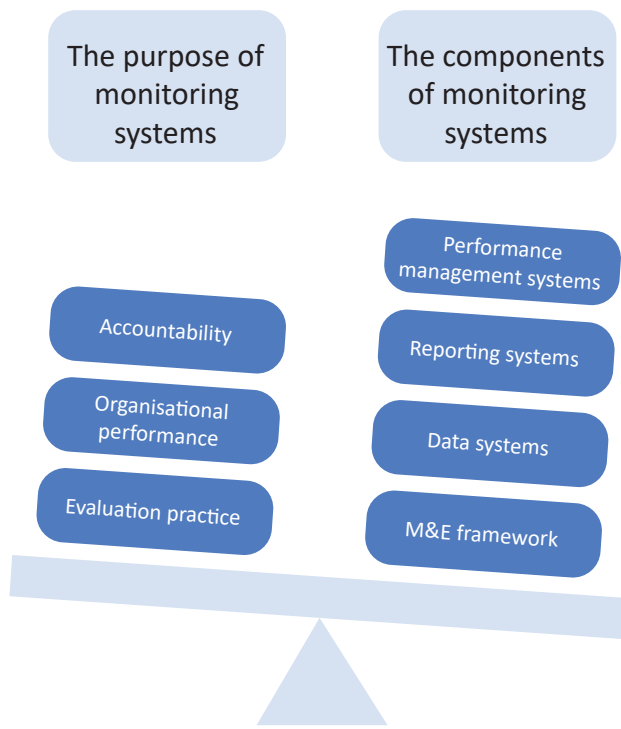
Figure 1.1: Components of monitoring systems

The figure above illustrates, on the left, three common reasons for monitoring: accountability, organisational performance, and evaluation practice. On the right, the column illustrates some components of a monitoring system: data systems, reporting systems, and performance management systems. While neither of these forms an exhaustive list, they are helpful anchors that can demonstrate how monitoring practice is not isolated and linear but rather links to a range of nested systems.

The following four sections explore the components of a monitoring system and their interconnectivity. Subsequently, a section is dedicated to each of the three purposes outlined in Figure 1. These sections examine the implications of a monitoring system's response to complexity and uncertainty, and discuss how these purposes impact the design of the monitoring system or its components.

The Components of a Monitoring System

One challenge in studying monitoring systems is the lack of widespread consensus regarding their definition, as organisations establish different boundaries around what constitutes monitoring. Some organisations fully integrate data and monitoring

systems, as well as reporting systems, while others maintain separate functional areas for monitoring and reporting. In certain cases, data management is entirely encompassed within the monitoring function, while in others, these areas have minimal connection. Similarly, some organisations integrate results, spending, and program implementation monitoring through performance-based budgeting. However, management of budget-related information is commonly done through separate channels, often with distinct management structures. Nonetheless, organisations must find ways to integrate these various components to facilitate effective monitoring.

The relationship among these components is multidirectional, encompassing different organisational dimensions. Accountability structures and reporting systems, for instance, incorporate a variety of policies and processes that contribute to accountability both within and outside the organisation. This includes accountability to donors and constituencies, beneficiaries, and achieving desired results. Additionally, accountability exists within the organisation, such as between an employee and their supervisor or between a programme and a monitoring and evaluation function.

Likewise, while data systems play a crucial role in providing inputs for monitoring practices, the outputs of monitoring activities also contribute to these same data streams. These circular dynamics emphasise that monitoring practice cannot be isolated from the complexity within which organisations operate.

While some data systems might be internal, focusing on process or financial data, others could be external, such as climatic or epidemiological data. This book questions one of the assumptions about monitoring, particularly in the chapters by Pophiwa and Everatt, which pertains to the location of monitoring functions. Traditional M&E assumptions propose that monitoring is an exclusively internal function, involving data generated and consumed by the organisation for programmatic decision-making. However, there are numerous situations where this is not the case. Citizen-based monitoring, parliamentary oversight as a form of monitoring, and the use of big data in monitoring are examples of crucial components of a monitoring system that exist outside an organisation. The locations of different components of a monitoring system may not strictly align with organisational boundaries, which can enhance the quality and use of monitoring. In an increasingly complex context, it may be necessary to draw upon data from various sources, thus altering the challenge of monitoring. For instance, organisational capacity to gather data becomes less of a concern when data is obtained externally. However, the processes of synthesising the data and tracing causal linkages differ when data originates from diverse sources. The implications of this remain relatively unexplored in the monitoring literature.

As global challenges become more complex and change occurs at a faster pace, monitoring data increasingly originates from diverse external sources. To effectively

utilise monitoring data for management processes, particularly when it is used for evaluation that connects results data to programmatic implementation data, a greater emphasis on causal linkages and attribution is necessary. In a context of complexity and uncertainty, adaptive management requires data that accurately reflects the activity's context. Results monitoring inherently captures the context through the gathered information, making it an essential tool for evaluation in an uncertain environment.

Monitoring for Evaluation and Learning

Monitoring of programmes establishes a foundation for comprehension of effectiveness, which occurs through two mechanisms. The first is process monitoring. Managers can determine whether programmes have been implemented as intended or not implemented at all, as process monitoring informs them about the implementation progress according to planned strategies. Achieving the intended results is unlikely if programs are not implemented correctly. The second mechanism involves monitoring the results. Some results necessitate causal inference to establish links with programmatic activities. Nevertheless, many programme results incorporate mechanisms for regular monitoring. For instance, conservation programmes benefit from continuous monitoring of biodiversity and land use. Similarly, health programmes may derive advantages from monitoring certain diseases or other health trends at a population level. While monitoring is often perceived to primarily focus on programmatic data and information, the monitoring of results, which may be conducted externally, also holds significant importance.

As monitoring data is primarily descriptive, the analysis occurs when users employ the monitoring data. The monitoring data tends to be devoid of inherent values, thus using and comprehending it requires an understanding of the context and the connections between the programme and its results. Various applications of evaluation data, such as strategic decision-making, adaptive management, and learning, are often associated with the evaluation function within an organisation.

The evaluation literature predominantly presents monitoring as an input for evaluation practice. Simultaneously, the literature acknowledges the significant constraints placed on the volume and extent of evaluation practice, with a disproportionate amount of organisational effort directed towards monitoring (Porter and Goldman 2013). A recent trend has emerged to broaden the definition of evaluation beyond its traditional practice, which typically involves occasional, externally commissioned activities. This expansion encompasses broader concepts of 'evaluative practice' and even includes 'evidence used in decision-making.' Expanding the terminology has profound implications for the relationship between monitoring and evaluation, as well as the role that monitoring can play in adaptive management and organisational learning.

From an evaluation standpoint, there is a logical flow of sequenced steps for designing a monitoring system: (i) a theory of change process, (ii) using this theory of change to develop evaluation questions, and (iii) aligning monitoring system components around these theory-based evaluation requirements (see Robertsen in this volume). A theory of change describes the way in which a programme's activities anticipate contributing to certain changes, which creates linkages between the activities implemented and the results achieved. Often, some of these causal linkages are very well established, while others are less certain. Monitoring can contribute to testing the accuracy of the theory of change, and can provide evidence to evaluate the process, the linkages to results, and the achievements of the results themselves.

From a monitoring perspective, there are several strengths and weaknesses to this approach. On the one hand, it is relatively straightforward to identify how monitoring data can be used. On the other hand, these theories of change address specific organisational needs that have developed in response to particular questions at a given time. Circumstances may change, as the rapidly changing context often reminds us. Theories of change are typically regarded as live documents and evaluation questions should adapt based on programmatic trends, although this is not always the case in practice. One advantage of monitoring is its ability to consistently track specific aspects over time. Establishing data collection tools, training staff, and reaching a consensus on common indicator definitions and their interpretations require significant investments of organisational time and resources. Regularly changing these elements would not only be inefficient but would also undermine the significant benefits of measuring core changes over an extended period. Maintaining consistency in indicator structure and other aspects of the monitoring systems enables managers to comprehend the results in response to programme or policy adjustments made to adapt to new circumstances.

In practice, monitoring systems are rarely designed to strictly align with the logic of evaluation needs, for better or worse. Typically, a negotiation process occurs between accountability requirements, reporting obligations related to the organisation's core operations, and evaluation needs. It is important to note that this negotiation process is not a zero-sum game. Monitoring data, which is necessary for accountability processes, can also contribute to organisational learning if there is a structured process for analysing and utilising this data for managerial decision-making. Moreover, an organisation's core business is not always fixed and can evolve over time through continuous reflection and discrete evaluation activities that propose significant changes. At this point, the monitoring components and evaluative components of M&E must reach a negotiated conclusion of how they can best complement each other.

Monitoring for Accountability

Significant research indicates that accountability is the primary motivation for monitoring, highlighting the importance of considering its implications for the design and use of a monitoring system (Bornstein 2006). An important implication, particularly in the context of Africa where a substantial portion of public sector activities rely on donor funding, is that organisations or departments may not have full autonomy in determining the key monitoring requirements (Goldman Byamugisha, Gounou, Smith, Ntakumba, Lubanga and Rot-Munstermann 2018). Many large bilateral or multilateral donor organisations enforce standardised systems of indicators and mandate all funding recipients to align with their bureaucratic monitoring priorities (Binnendijk 2019). This raises critical questions regarding the organisational outcomes that the investment in monitoring aims to enhance. Moreover, this issue is closely intertwined with the neocolonial origins of monitoring and evaluation practices across Africa, which have perpetuated entrenched donor interests (Mapitsa, Tirivanhu and Pophiwa 2019).

Many organisations receiving donor funding are experts at aligning organisational needs with donor needs to reduce the burden of coordination. Nonetheless, it undoubtedly demands organisational resources to align various programme logics and ensure that data gathered for a single purpose can serve several others. A second implication is that incentives to game the data often arise if monitoring data is primarily used for accountability to stakeholders rather than the intended beneficiaries. In this volume, Masvaure considers the dependency of these incentives on the way in which monitoring results are coupled with consequences.

The primary use of monitoring for accountability creates a blanket incentive to overstate the extent of implementation and performance quality, which limits the extent to which monitoring can foster learning. The implementation of various strategies can reduce the ease at which monitoring results can be gamed (Mizrahi 2017). However, perhaps the least complicated and most effective strategy involves shifting the use away from an exclusive focus on accountability. Studies show that public officials view using performance information for accountability as counterproductive and ineffective (Behn 2003; Heinrich and Marschke 2010; Moynihan 2008). This results in low investment in the assurance of high quality data, as well as compliance with monitoring requirements.

In public sector contexts in Africa, where the demand for accountability by citizens, donors, parliaments, and other role players is legitimately high, using monitoring data as a means to provide accountability is an attractive solution.

However, it neglects to acknowledge the context of complexity and the presence of multiple embedded systems that exert influence on monitoring practice. It also

fails to recognise the existing body of research that delves into the practical aspects of monitoring. The accountability function of monitoring establishes an oppositional relationship between the generators and users of monitoring data, thereby creating incentives for subpar data quality. Complexity, as a contextual factor, not only influences monitoring data but can also be utilised to challenge the interpretation of such data. Scholars studying evidence use in decision-making extensively discuss this phenomenon (Stewart, Dayal, Langer and van Rooyen 2019). However, most approaches to monitoring systems assume that organisations provide a framework for interpreting and analysing monitoring data for decision-making. While this assumption may hold true in certain cases, theories of change processes often highlight that achieving this level of consensus cannot be taken for granted.

Organisations make significant investments to ensure that the monitoring data used for accountability remains untainted by incentives that could compromise its integrity. Chapter 6 of this volume explores algorithmic governance, an extreme form of coupling data and decisions. However, a more prevalent practice involves linking decision-making to processes such as parliamentary oversight, systems for data quality checks, or imposing sanctions for attempts to manipulate the data. While evidence suggests that some of these processes are effective, they also consume a considerable amount of time, thereby limiting the utility of the same dataset for adaptive management or responding to an uncertain context. The design of the components within a monitoring system highlights the trade-offs between these different objectives, making it challenging to repurpose the generated monitoring data from one objective to another. Occasionally, there are shared attributes of monitoring systems that effectively serve both accountability and adaptive management purposes. These attributes include timeliness, accuracy, high-quality data, and completeness. However, in other cases, as the investment in accountability systems grows, the system becomes less suitable for adaptive management purposes. These aspects of both mutual reinforcement and trade-offs are discussed throughout the volume in various ways.

Monitoring to Improve Organisational Performance

First and foremost, organisations use monitoring as a tool to improve their performance. The underlying rationale is that organisations can gather sufficient information to optimise their practices and outcomes by continuously tracking actions and achievements at higher levels of the results chain. When implemented effectively, monitoring can complement the evaluation function by providing focused insights in areas where ongoing data collection cannot be addressed due to complexity, uncertainty, or specific research gaps in the contextual landscape.

Several factors contribute to the usefulness of monitoring as a tool. Its continuous nature enables ongoing tracking over time, which can inform adaptive management practices (Waylen Blackstock, Van Hulst, Damian, Horváth, Johnson and Van Uytvanck 2019). This aspect is particularly valuable during periods of uncertainty and when dealing with high levels of complexity. While evaluations can address critical questions regarding program design, they cannot provide the same continuous feedback required for adapting to changing contexts.

Secondly, monitoring data offers insights into the implementation of organisational activities, allowing for triangulation with various other data sources. These sources may include internal organisational data such as performance-based budgeting, data gathered during evaluation processes, national statistical data, or relevant demographic trends. Organisations can gain a comprehensive understanding of their operations by analysing trends and linkages between their impact and interconnected factors. This process is particularly valuable given the increasingly complex nature of social problems and the involvement of multiple stakeholders.

The use of monitoring data for performance management implies a few things. It is essential for monitoring data to be built into several systems that measure results, negotiate the outcomes of these measurements, and reflect this to people at different places within the organisation. People feel motivated to feed into an effective system upon seeing that their efforts yield results. The feedback loop regarding processes of use is often broken (Mapitsa, Tirivanhu and Pophiwa 2019). An effective monitoring system for improving organisational performance should ensure that processes are in place to feed back data use to those contributing the data. A robust monitoring system allows each contributor to understand how their part of the data puzzle fits into a coherent whole. Simultaneously, coherence and management of these various processes is necessary. In an organisational context characterised by uncertainty, feedback time is of the essence; the loop from data to decisions to feedback needs to be clear and timely. The loop must be adaptive enough that necessary changes can be made for efficient implementation in a context of complexity, should anything change.

Examples of Monitoring in Complex, Uncertain Systems

The section below presents examples of programmatic interventions that took place in complex or uncertain contexts and examines the ways in which monitoring systems either contributed to effective programme design, or failed to do so due to inappropriate institutionalisation or barriers hindering the use of monitoring data.

Policymakers in the COVID-19 pandemic faced a situation where they had relatively good access to external data about the pandemic results. This data included the number of deaths each day, the rates at which people tested positive for COVID-19,

the availability of hospital beds, and similar information. The health sector had widely available information systems, although the quality of this data and its causal mechanisms could be questionable. They eventually obtained information about the virus, such as its usual rate of spread and the conditions under which it spread. However, the challenge they faced was building causal linkages between those pandemic results and the programmatic or policy mechanisms at their disposal. Tools to combat the pandemic, such as mask mandates, curfews, and limits on gatherings and mobility, were connected to pandemic outcomes. However, due to the rapid rate of change of the pandemic, the differing contexts in which the policy response was rolled out, variations of policy implementation and enforcement, and the comparably long timeline for rigorous evidence generation, determining the causal linkages between policy decisions and pandemic outcomes has been challenging. For example, some analysts believed the lockdown was relatively successful in 'flattening the curve' of infections, while others believed it led to spikes in infection in areas with townships of high population densities and little possibility for social distancing. The contestation of evidence regarding causal linkages has shown that even when high-quality monitoring data is relatively widely available, disputes about the interpretation of this data arise from ideological positions and methodological preferences. While robust monitoring systems can shape the scope of the debate around causal linkages, they cannot eliminate contestation.

Monitoring systems shaped decision-making in the following example: A critically endangered member of the Lily family grew exclusively in a small provincial reserve in Limpopo. There had been a few censuses of the plant, but the biological knowledge of the species was relatively low. Conservation decisions needed to be based on good practice in general protected area management. Although the reserve held ecological importance, it lacked a tourism market or other economic benefits. Moreover, it did not receive significant local support as it was seen as a means to exclude people from accessing the land. Despite breaching the allowed land use, cattle grazed widely. A small lobby group of environmental activists and scientists, concerned with this particular protected species, organised a collaborative intervention to strengthen land management within the provincial reserve for the benefit of the plant. They worked alongside local and provincial authorities to reinforce fencing, garner support for the conservation effort, and prevent further grazing within the reserve. Their intervention was framed based on indicators for protected area management. Process monitoring was established to track improvements in protected area management practices. The implementation process of the conservation initiative was largely successful, as grazing decreased due to the newly erected fences. However, results monitoring occurred only at a later stage, when an additional count of the lily's population was conducted. It became apparent through this count that the interventions had significantly reduced the size of the lily's population. Grazing was found to be essential in reducing competition faced by the plant, as faster-growing

grass crowded it out. The decisions made based on the best available evidence did not account for all components of this complex system. This example illustrates that effective monitoring systems have the potential to challenge established theories of change and programmatic designs.

The establishment of the social relief of distress grant serves as a final example involving the response to the COVID-19 pandemic. This grant expanded social grants available in South Africa to unemployed adults who would not qualify for other grants, such as the old age or disability grant. Since the beginning of the lockdown and subsequent years, there has been contestation and debate surrounding this grant, highlighting divergent views regarding its purpose and the necessary monitoring to determine its effectiveness. Evidence suggests that it played a significant role in reducing hunger. However, several other intended benefits, such as facilitating job seeking, remain subjects of significant contention. The data required to monitor its effectiveness is spread across multiple departments that lack aligned systems for generating useful monitoring data on key aspects of the intervention. Moreover, numerous variables have compounded the desired outcomes of the grant, including poverty reduction, increased employment, and improved access to food. The dynamic nature of the pandemic has made it challenging to isolate the contribution of this particular intervention.

Enablers and Constraints to Effective Monitoring

Many organisational factors such as a lack of awareness and the resulting deprioritisation, budget and mandate limitations, hesitancy around transparency, or limited ability to implement changes based on available data, can hinder effective monitoring. However, monitoring systems have been strengthened due to significant and rapid change in the contextual factors that previously impeded progress.

The COVID-19 pandemic has highlighted the importance of data in policy decisions. The availability of information regarding crucial aspects of the pandemic response, including hospital bed availability, virus spread rates, recovery rates, and vaccination rates, has driven the response of many countries and undoubtedly saved many lives. Paired with the best available knowledge about behavioural science and decision-making, the pandemic has inspired transformative changes at a societal level that would have been unimaginable before. On one hand, this underscores the heightened level of uncertainty in which monitoring systems now operate. The previously assumed impossibility of many societal changes now taken for granted prompts questioning of numerous other issues.

Citizen expectations regarding the availability and use of evidence have shifted, evident in various fields. The expectation now includes real-time tracking of social assistance measures, mirroring the accessible information on the pandemic, owing

to increased awareness and involvement with diverse data sources. The government COVID-19 portal has become a household decision-making tool, resulting in profound implications for expectations concerning democratic decision-making and transparency.

Simultaneously, false news has showcased the extent to which data can be contested and interpreted differently by various stakeholders (Uwalaka, Nwala and Chinedu 2021). This diverse interpretation emphasises the significance of monitoring systems that not only provide data but also establish a framework for its interpretation and use.

Interests and incentives linked to the interpretation of monitoring data often lead to ongoing contestation of monitoring systems. Nonetheless, comprehending the origins of this contestation and its implications for monitoring system design is crucial to ensure their suitability for the intended purpose.

Conclusion

The government needs to respond to increasingly complex and unpredictable challenges, highlighting the heightened importance of monitoring systems. However, these systems must be capable of measuring what is crucial for adaptive management. It is essential to have the capacity to (i) collect the requisite data, (ii) analyse it, and (iii) provide feedback on its use to those responsible for gathering it. The organisational systems required for these tasks are intricate and must operate with timeliness to be effective. Governments already invest substantial time, energy, and resources into monitoring, often focusing on compliance and accountability, which can compromise the quality of the gathered data. While this serves a purpose, it also carries risks that can incentivise manipulation or improper use of monitoring data. There are indications that monitoring systems could be better structured to align with performance improvement objectives.

The unprecedented demand for transparent and readily available information, combined with existing capabilities, presents significant opportunities for building monitoring systems that are tailored to their purpose and fulfil both information supply and demand. With big data and technology accessible to virtually all citizens, who can provide real-time verifications or input on implementation and delivery, the potential for creating monitoring systems that meet the necessary requirements is substantial.

Reference List

Basheka, B. C. & Byamugisha, A. K. 2015. The state of monitoring and evaluation (M&E) as a discipline in Africa: From infancy to adulthood?

Behn, R. D. 2003. Why measure performance? Different purposes require different measures. *Public Administration Review*, 63, pp. 586–606. https://doi.org/10.1111/1540-6210.00322

Binnendijk, A. L. 2019. Donor agency experience with the monitoring and evaluation of development projects. *Methods for social analysis in developing countries.* pp. 165-184. https://doi.org/10.4324/9780429043239-10

Booth, D. & Lucas, H. 2000. *Good practice in the development of PRSP indicators and monitoring systems.* London: Overseas Development Institute.

Bornstein, L. 2006. Systems of accountability, webs of deceit? Monitoring and evaluation in South African NGOs. *Development*, 49(2), pp. 52-61. https://doi.org/10.1057/palgrave.development.1100261

Goldman, I., Byamugisha, A., Gounou, A., Smith, L. R., Ntakumba, S., Lubanga, T. & Rot-Munstermann, K. 2018. The emergence of government evaluation systems in Africa: The case of Benin, Uganda and South Africa. *African Evaluation Journal*, 6(1), pp. 1-11. https://doi.org/10.4102/aej.v6i1.253

Heinrich, C. J. & Marschke, G. 2010. Incentives and their dynamics in public sector performance management systems. *Journal of Policy Analysis and Management*, 29 (1), pp. 183–208. https://doi.org/10.1002/pam.20484

Kusek, J. Z. & Rist, R. C. 2004. *Ten steps to a results-based monitoring and evaluation system: a handbook for development practitioners.* Washington, DC: World Bank.

Mapitsa, C. B. & Korth, M. T. 2017. Designing diagnostics in complexity: Measuring technical and contextual aspects in monitoring and evaluation systems. *African Evaluation Journal*, 5(1). https://doi.org/10.4102/aej.v5i1.196

Mapitsa, C. B., Tirivanhu, P. & Pophiwa, N. Eds. 2019. *Evaluation Landscape in Africa.* African Sun Media.

Mayne, J. & Zapico-Goñi, E. 2017. Effective performance monitoring: a necessary condition for public sector reform. In *monitoring performance in the public sector*, pp. 3-30. Routledge. https://doi.org/10.4324/9781315124681-2

Mihaiu, D. 2014. Measuring Performance in the Public Sector: Between necessity and difficulty. *Studies in Business & Economics*, 9(2).

Mizrahi, S. 2017. *Public policy and performance management in democratic systems: Theory and practice.* Springer. https://doi.org/10.1007/978-3-319-52350-7

Moynihan, D. P. 2008. *The dynamics of performance management constructing information and reform.* Washington, DC: Georgetown University Press. https://doi.org/10.1353/book13015

Nalubega, T. & Uwizeyimana, D. E. 2019. Public sector monitoring and evaluation in the Fourth Industrial Revolution: Implications for Africa. *Africa's Public Service Delivery and Performance Review*, 7(1), pp. 1-12. https://doi.org/10.4102/apsdpr.v7i1.318

Porter, S. & Goldman, I. 2013. A growing demand for monitoring and evaluation in Africa. *African Evaluation Journal*, 1(1), p. 9. https://doi.org/10.4102/aej.v1i1.25

Stewart, R., Dayal, H., Langer, L. & van Rooyen, C. 2019. The evidence ecosystem in South Africa: growing resilience and institutionalisation of evidence use. *Palgrave Communications*, 5(1). https://doi.org/10.1057/s41599-019-0303-0

Uwalaka, T., Nwala, B. & Chinedu, A. C. 2021. Social media, fake news and fake COVID-19 cures in Nigeria. *Journal of African Media Studies*, *13*(3), pp. 435–449. https://doi.org/10.1386/jams_00058_1

Vigoda-Gadot, E. & S. Mizrahi. *Managing Democracies in Turbulent Times.* Springer-Verlag Berlin AN, 2016.

Waylen, K. A., Blackstock, K. L., Van Hulst, F. J., Damian, C., Horváth, F., Johnson, R. K. & Van Uytvanck, J. 2019. Policy-driven monitoring and evaluation: Does it support adaptive management of socio-ecological systems? *Science of the total environment*, *662*, pp. 373–384. https://doi.org/10.1016/j.scitotenv.2018.12.462

Wholey, J. S. 2007. *Monitoring performance in the public sector: Future directions from international experience.* Transaction Publishers.

Yung, L., Louder, E., Gallagher, L. A., Jones, K. & Wyborn, C. 2019. How methods for navigating uncertainty connect science and policy at the water-energy-food nexus. *Frontiers in Environmental Science*, *7*, p. 37. https://doi.org/10.3389/fenvs.2019.00037

Chapter 1

Chapter 2: Strengthening Equity-Informed Monitoring Systems Within the Public Service

Philip Browne

Introduction

Governments worldwide implement results-based performance monitoring frameworks to measure and evaluate the programme progress and effectiveness, aiming to enhance performance and achieve strategic development priorities. These frameworks serve two interconnected and mutually reinforcing functions of good governance. Firstly, they ensure accountability by assessing state performance in relation to expenditure, guaranteeing that government services, funded by taxpayers, are accessible and equitable to all citizens. The routine M&E systems employed by governments across sectors tend to prioritise the compliance and accountability aspects of good governance, aligning with positivist requirements.

However, this chapter asserts that the other facet of good governance, namely the imperative to comprehend why significant portions of society remain underserved and marginalised, demands greater attention. The summary presented in Figure 1 illustrates that equity and inclusion must be integral characteristics of good governance (Agrawal Kalugampitiya, Rinxin & Hashim 2017).

Figure 2.1: *Agrawal et al, The Place of Equity and Inclusion in Good Governance Systems. 2017*

In the context of public sectors, good governance encompasses the structures, systems, and government entities that collaborate to facilitate evidence-informed decision-making processes and the execution of policies, programmes, and projects aimed at fostering equitable economic and social development, ensuring that no one is excluded. At the core of public sector good governance lies the monitoring and evaluation process, which serves as a mechanism for performance management and accountability.

Within this expansive domain of good governance, there exist opportunities for leaders and practitioners in monitoring and evaluation to critically examine their practices and identify approaches to enhance monitoring systems with a focus on equity.

The first step involves mapping out the theoretical underpinnings of equity-informed monitoring, evaluation, and learning practice and explaining why this approach

goes beyond compliance to serve a more developmental purpose. This chapter will unpack some of these theoretical positions on the role of equity in M&E and examine how these understandings are being translated into emerging technical practices in different sectors. It will also address the types of capacities that need to be nurtured to achieve equity-oriented monitoring.

The public sector should question whether evaluation, as a practice, contributes to entrenching and maintaining conditions of inequality or if it holds the potential to act as a transformative developmental agent through equity-driven evaluation methodologies. Morkel (2021) argues that evaluation for transformational change or transformative evaluation should be grounded in social justice and equity in Africa. In her blog post, Jara Dean-Coffey, a specialist and writer in the field of evaluation, emphasises that "our distance from the work/people/issue/community has often led to us not seeing/feeling what is happening, but also not understanding the nuance and complexity that exists in the human experience, let alone the planet or universe".[2] The African Evaluation Association (AfrEA) acknowledges this omission and urges evaluators to consider that "evaluation considers issues and norms that are sensitive and important in African contexts, including power dynamics, the relationships between people, the policies and priorities for development, different ideas about what 'success' is and how it can be measured, and the balance between the rights of individuals, societies and nature".[3]

The rigours and constraints imposed on government-led M&E processes are inherently technocratic rather than people-centred. Consequently, M&E processes are often described as mandatory, compliance-focused, extractive, burdensome, and costly. This chapter argues that such terms would be less frequently used if conventional M&E methodologies became more inclusive, equity-driven, and participatory. This necessitates transforming the discourse surrounding the conceptualisation and utilisation of M&E by practitioners, moving beyond mere data collection and information processing, and embracing a comprehensive commitment to equity and inclusion. In the African evaluation context, evaluators must recognise that development requires a complex systems-informed evaluation approach that integrates sociocultural, economic, political, technological, and environmental factors, connecting local contexts with the global stage and Africa with the world (AfrEA).

Many reasons drive the instinct to actively improve and strengthen government monitoring and evaluation systems. Fiscal pressures and the increasing expectations of ordinary citizens continuously motivate governments to expand services and

---

2   Jara Dean-Coffey. 2021. Positionality posted on the Equitable Evaluation Initiative website at https://www.equitableeval.org/blog-main/positionality

3   African Evaluation Association. 2021. The African Evaluation Principles at https://afrea. org/AEP/new/The-African-Evaluation-Principles.pdf

enhance quality standards (Mackay 2007). This includes addressing the expectations from vulnerable, marginalised, and underserved citizens in an equitable manner. However, achieving this may be challenging as vulnerability and marginalisation within societies often stem from structural inequities, encompassing political, economic, and social dimensions. Without understanding and measuring these factors comprehensively, governments face difficulties in effectively addressing them. Additionally, when dealing with broad concepts such as poverty, inequality, and marginalisation, the diverse and nuanced aspects of people's lived experiences are often overlooked. Generalised, insufficient, outdated, or poorly disaggregated data, as well as data gaps, can lead to service gaps, leaving segments of the population unaccounted for in national statistics. Consequently, the lack of evidence negatively impacts the state's ability to meet the needs of the most vulnerable and marginalised groups in society, weakening its capacity for equitable service delivery.

The Changing Landscape of Government Monitoring Systems

Practitioners generally face the challenge of being unaware of critical equity and inclusion-related data. Holmes (2020:1) highlights the concept of "positionality", which refers to an individual's worldview and the position they adopt regarding a research task and its social and political context. The concept of positionality challenges the notion that the practice of M&E should be objective and free from values. Consider two government M&E officials working with police crime data and engaging with statistics related to gender-based violence (GBV), including rape, sexual assault, and femicide. Each official brings their professional and personal perspectives to the data. One may possess knowledge, empathy, and responsiveness to gender-based inequalities, while the other may hold a patriarchal understanding of gender roles. However, these perspectives can converge when acknowledging that each data point represents a complex individual with unique circumstances in their lived reality. This understanding allows these practitioners to shape the evidence to inform equity-responsive analysis effectively.

Equity-driven monitoring systems require practitioners who understand and are committed to using evidence to measure both the drivers and consequences of inequalities. Achieving this necessitates capturing evidence that reveals structural issues (drivers) and the socioeconomic impacts (consequences) of inequalities within societies, considering how they affect different segments of the population. In the case of police data on GBV, combining disaggregated data sets can provide a more detailed understanding of the victims and survivors, encompassing factors such as age, ethnicity, geographical location, marital status, and education level to a broader degree. At a more granular level, additional information such as social and economic

status, level of education, sexuality, and HIV status can be considered.[4] Working with a more comprehensive cross-section of data allows practitioners to identify patterns and intersections of inequity and exclusion. This, in turn, enhances their ability to provide evidence that informs the work of policymakers and programme designers. The Organisation for Economic Cooperation and Development (OECD 2018:72) emphasises the importance of national statistical systems pursuing sophisticated data disaggregation strategies. Current statistics typically capture national averages but fail to reveal disparities at the subnational, community, household, and individual levels. AfrEA advocates for evaluations that serve an equity purpose, recognising their potential as transformative exercises. They further emphasise that financing, commissioning, conducting, and utilising evaluations in Africa are "highly responsible tasks, especially when dealing with vulnerable communities and economies, developing institutions, and the rich diversity of worldviews, experiences, and traditions that define African societies" (AfrEA).

Monitoring systems that are designed to produce equity-informed evidence are most effectively managed by practitioners who comprehend the importance of moving beyond routine or standardised measurement processes and utilising the broader potentiality of data. The pervasive challenge of poverty in Africa and the way in which it is measured can be examined. In the development context, poverty is generally seen as the overarching manifestation of inequality. This is highlighted by the fact that poverty and inequality feature prominently in the United Nations' seventeen Sustainable Development Goals (SDGs).[5] In the broadest terms, the most widely used measure of inequality is the Gini coefficient, which ranges from zero (perfect equality) to one (perfect inequality). The Gini Coefficient is a generalised, population-based metric. It is not easily broken down to reveal the sources and consequences or the sections of the population most affected by inequality. Inequalities and disadvantages are embedded in the power dynamics of social structures such as class systems, cultures, religions, and gender relations, and often manifest in social institutions and socioeconomic systems. A more composite and multidimensional response is required to better understand the complexity and multidimensionality of the drivers of poverty and inequality, given that narrowing gaps in one area may not be sufficient to reduce disparities in other domains of well-being (Rohwerder 2016). Statistics South Africa (StatsSA) has undertaken work on a nuanced approach to the consideration of the multidimensional and intersectional forms of poverty to

---

4   The concept of "granularity" in M&E refers to the level of detail of data, based on the understanding that the more that detailed that data is, the more precise any analysis can be.

5   The SDGs were adopted by the United Nations in 2015 as a universal call to action to end poverty, protect the planet, and ensure that by 2030 all people enjoy peace and prosperity – see https://www.undp.org/sustainable-development-goals

develop a multidimensional poverty index, intended to move away from the blunt measurement of poverty, such as the World Bank's $1.90 a day threshold.[6]

An added layer of complexity involves the descriptors associated with the concept of inequality, such as vulnerability, marginalisation, disadvantage, deprivation, disempowerment, and social exclusion. Bok (2018) argues that social exclusion can be described as a comprehensive, multidimensional, and dynamic concept that generally refers to the limited opportunities of individuals to participate financially, socially, culturally, and politically in their societies. Khan et al. (2015) suggest that social exclusion is a process by which certain groups are systematically disadvantaged due to discrimination based on their ethnicity, race, religion, sexual orientation, caste, descent, gender, age, disability, HIV status, migrant status, or area of residence. It further refers to the processes behind the accumulated vulnerability and weakening of social rights, including discrimination embedded in public institutions, such as the legal system or education and health services, and social institutions like religion, cultures, workplaces, family networks, and the household. Sen, Kessler and Loveridge (2018) proposes a capabilities approach and argues that social arrangements should be evaluated according to the extent of freedom with which people can promote or achieve functionings that they value, suggesting that wellbeing should be measured according to what individuals can do (capabilities) as opposed to what they do (functionings). This poses a challenge for M&E practitioners who need to understand the assumptions made by Sen et al. (2018), and the way in which to develop tools that can measure capabilities.

Applying an Equity Lens to Monitoring Systems

An equity perspective makes the production and use of data fairer, more robust, and more accurate in general. Moreover, to ensure equity in any analysis process, the data being used must reflect the fact that an individual's experiences are not unidimensional (poor, illiterate, or HIV positive) but are based on multiple and intersecting dimensions, identities, and experiences. As an example, the effects of poverty, marginalisation, disability, and sexuality would intersect within the reality of a person who is a poor, rural, disabled lesbian of ethnic minority. Different combinations of demographics and identities create different types and experiences of inequality for various people. From an equity point of view, it is essential to gear monitoring systems toward generating data and analysis that reflects these very granular realities. Population-level data for disability (2.6% of the population) does not reflect different kinds or experiences of disability. Neither does it reflect the different levels of opportunity and access that disabled persons from different backgrounds may have. When monitoring systems work with blunt data, they tend to overlook or conceal the

---

6   The World Bank data portal at https://data.worldbank.org/indicator/SI.POV.DDAY

different socioeconomic realities of population sub-groups and, consequently, reduce analysts' ability to dissect issues of diversity, intersectionality, and difference. This lack of multidimensional data could impact policymaking where policies are designed based on inadequate or insufficiently differentiated evidence.

Figure 2.2:   *Wheel of intersectionality, Identiversity, [1 July, 2023]*

Common drivers of inequality relate to a lack of inclusive growth, lack of investments in human capital, lack of pro-poor fiscal policies and redistribution, lack of access to essential services and human rights, and a lack of political will to tackle the root causes of discrimination, as well as structural exclusions based on social, political, cultural, and economic factors. Every country has its unique systemic challenges in this regard, but these dynamics often play a role in perpetuating social, economic, and political inequities within societies. Issues related to legal discrimination, social expectations regarding gender and sexual identity roles, restrictions on bodily integrity, class and ethnic forms of discrimination, and various forms of exploitation and marginalisation, largely explain the persistence of unequal outcomes in employment, entrepreneurship, health and wellbeing, access to opportunities, and political representation.

With women making up just over 50% of the world population, gender is perhaps the most contentious area of inequality. The Organisation for Economic Cooperation and Development (OECD) notes that discrimination in social institutions, such as education, health provision, the labour force, and financial access, contributes to

obstacles in gender equality in development outcomes across all world regions.[7] However, the commitment to the implementation of equity-based monitoring and evaluation systems remains rhetorical, unless M&E systems within public services are able to develop and implement indicators that measure progress on achieving gender equality by understanding that the lived experience of women and girls is complex, heterogenous, and multidimensional.

The term 'institutional', as referred to by the OECD, speaks to the formal and informal systems, rules, and norms that structure and govern the social order and that obstruct or exclude people from social service provisioning, public employment, or other areas of social interaction (Fischer 2011). The parameters for acceptable decisions, choices, and behaviour for women and other marginalised groups in society are established by social institutions, and consequently define their roles and impact their life outcomes. These are undoubtedly challenging areas to measure if conventional M&E tools are used. Religious and cultural norms regarding sexual and reproductive health typically oppose notions of equity and rights, resulting in the disruption of equity-informed evidence for M&E practitioners to work with. The data challenges associated with an equity-oriented approach to managing data on abortion, contraception, age of consent, sex trafficking, cyber-grooming of underage girls, and female genital mutilation, should be considered. Gender-based discrimination in rights, opportunities, and outcomes interconnect and overlap, thereby further reinforcing women and girls' marginalisation.

Discriminatory laws, norms, and practices as measured by the OECD's Social Institutions and Gender Index (SIGI) and the African Development Bank's (AfDB) Africa Gender Index (AGI), map the limitations placed on the role of women as sexual beings, workers, entrepreneurs, healthy citizens, leaders at the national and local levels, and actors of human development. Based on these kinds of indices, higher levels of gender-based discrimination in social institutions are associated with lower equality in outcomes. Simplified, this means that countries with higher levels of discrimination in social institutions are further from achieving gender parity. In its 2020 report, the AfDB notes with emphasis that the limited availability of policy-relevant gender statistics poses a challenge to the inclusion of several fundamental aspects of gender inequality in the AGI.[8] The recognition that national and regional statistics are lacking in critical areas where exclusion and marginalisation are most pronounced is a positive step forward for advocates of a more equity-driven monitoring practice that integrates this approach into monitoring systems.

Fisher (2011) argues that exclusion is a pressing development concern. An equity-driven monitoring approach highlights the intersecting processes of exclusion which

7   OECD at https://www.oecd.org/dev/development-posts-gender-discrimination.htm
8   Africa Gender Index: Methodological and Statistical Report 2020:7.

are not effectively captured by poverty, inequality evidence, and methods of analysis, particularly those within contexts of high or rising inequality. Absolute and relative indicators often reveal little about processes of exclusion and marginalisation, and if they do, they tend to do so by revealing clues about the spaces within which exclusionary practices and processes might operate. It is understood that poverty is a problem statistically, but the various ways in which different "poor" or "margina-lised" people are impacted is not as clear. Standard statistical sampling methods based on generic outcome indicators are the customary practice in the government and development sectors but are poorly suited for approaches that focus on equity or social inclusion. An example would be the way in which a blunt statistic, such as the HIV+ percentage of a population (usually disaggregated by age and gender), fails to capture differential experiences of access to testing, treatment, care, and support, as well as community and internalised stigma. To this end, more inductive methods, such as gathering granular data at the ground level, would better serve monitoring systems that can trace the kinds of implicit and explicit socioeconomic dynamics that affect people's lives. This would include interdisciplinary analyses of structural and institutional disjunctures and asymmetries operating across social hierarchies and among comparable cohorts within a social hierarchy, such as those with similar levels of educational achievements and employment expectations.

Reimagining the Monitoring Paradigm

The utilisation of the tools of equity, inclusion, and intersectionality requires a shift in methodological approaches to data collection and analysis and an epistemological shift in terms of what constitutes valid and recognised data. Part of the epistemological shift has been the movement from hard, empirical (quantitative) data to a growing interest in the more qualitative dimensions of intersectionality, which form an integral analytical layer for a better understanding of the complexity of inequalities and social exclusion. Originally posited by legal scholar Kimberlé Crenshaw in the field of critical legal studies, the concept of intersectionality aimed to deconstruct the application of laws utilised in legal cases to illustrate the ways in which the structures of law and society could be intrinsically racist. Crenshaw (1989:167) noted that intersectionality is a strategic way to place those who are currently marginali-sed at the centre of the conversation and is "the most effective way to resist efforts to compartmentalise experiences and undermine potential collective action". The concept of intersectionality as a framing construct has subsequently been applied in various disciplines, including radical sociology, feminism, gender studies, queer theory, equity and diversity studies, and critical race theory.

The concept of intersectionality has become a contentious term in the lexicon of conservative politics for its work in deconstructing racist, colonial, sexist, and homo-phobic discourse. Despite resistance, the concept of intersectionality has become a

valuable analytical tool throughout the last decade, particularly for those working on issues related to socioeconomic inequalities and marginalised or vulnerable populations within societies. While it has been broadly applied in qualitative research studies, it has only recently experienced a surge in quantitative research, owing to the technical and cost challenges associated with collecting more granular quantitative evidence. However, the implementation of a theoretical intersectionality framework into quantitative data analyses is gaining increasing interest in health research due to the understanding that the complex causes and mechanisms leading to health inequalities can be improved by the integration of an intersectionality framework (Mena et al. 2019). This can be similarly applied to other social sectors such as education, criminal justice, and social protection, where inequities constitute a significant determinant of access to fair and appropriate services. It is evident that people may continue to 'fall through the cracks' in policymaking and service delivery due to the lack of data to evidence and understand the unique challenges experienced by people facing multiple forms of inequality.

Many countries have constitutions that enable states to prioritise human rights, inclusion, and equity for all citizens, and rights-based pieces of legislation and policy frameworks that stem from crucial constitutional mandates. Some countries retain policies and legislation that undermine efforts to achieve equity and inclusion despite having a rights-based constitutional mandate. Challenges arise when government sectors are expected to realise different kinds of equity imperatives in their work or services, with these being dependent on existing policies as well as the quality of data available to policymakers and planners. In budget-constrained countries where public services are often inadequate, many citizens slip through the inclusion net due to their specific vulnerabilities not being captured in official data sets. Individuals who are at the intersections of disadvantages may struggle to have their needs met when "policies are developed using a single-factor lens, activated by single-factor trigger points, and/or developed to offer single-factor interventions" (Corus et al. 2016); they are "invisible" to official recognition. Crenshaw, and other feminists, point to the intersection of race and gender in countries such as the United States and Brazil, where women may face exclusion from jobs deemed more appropriate for men due to their sex. Women may be excluded from jobs considered "women's jobs" because of their race (AWID 2004). As a result, women of ethnic minorities specifically face exclusion from employment opportunities. If official data collection and analysis processes do not recognise these experiences, there is little likelihood that they will be addressed in any viable policy or programmatic way. It is evidenced by these methodological challenges that there is enormous scope for M&E and knowledge management practitioners, who work at the fulcrum between data generation and data analysis, to interrogate their practice and start applying new methodologies and tools to sharpen an equity-informed intersectional practice.

There exists a tendency within government to foreground the importance of stakeholder consultation and community engagements to foster understanding regarding people's lived realities, but these engagements can be formulaic and are incapable of allowing communities to feed local experience and knowledge into government monitoring systems. This highlights a lost opportunity, as government responsibilities (including the provision of healthcare, education, and social assistance) require ever more sophisticated evidence sources to enable the design of policies and programmes that are as inclusive as possible and are informed by the actual needs of people. The tendency of people being omitted from data is particularly distressing, as government policies and programmes that produce administrative data can inadvertently contribute directly to creating, enabling, and sustaining institutional and structural forms of discrimination and marginalisation. The desirable condition is one that allows cross-sector data sharing and integration, which enables the transformation of individual-level information into actionable intelligence for M&E practitioners. Ideally, more granular evidence can be used to understand urgent and long-term community needs and can result in the implementation of improved and more equitable services, systems, and practices that, according to the UN concept, "leave no one behind".

Reflections on the Discourse of M&E

Considering the reflections on what could be considered a "desirable state" for an equity-informed M&E practice, it is crucial to acknowledge the ideological challenge posed in terms of the implementation of an equity-informed monitoring process. Concepts such as human rights, social inclusion, and equity are socially constructed and hinge on a consensus regarding the definition of 'normal' in each society. As an example, 2021 saw 31 countries throughout the world legally recognise same-sex marriage, while 71 countries criminalise same-sex sexual activity, and 11 of those countries impose the death penalty in response to it.[9] Defining the concept of 'normal', 'acceptable', or 'decent' is not a simple task in many developing countries, where many people do not enjoy what may be considered a decent standard of living, due to the perspectives that countries or regions may have on critical developmental challenges. People who find themselves outside of their specific societal definition of 'normal' may be excluded based on their race, caste, or gender, due to social exclusion being structured around hierarchy. It is often difficult to 'objectively' identify those being socially excluded, as it is a matter of assumptions made, criteria adopted, and judgements used, which results in the concept of social inclusion and exclusion being contested. Equity and inclusion concepts focus attention on two central elements of deprivation: its multidimensionality and the processes and social

---

9   Human Rights Campaign. 2022. Marriage Equality Around the World https://www.hrc.org/resources/marriage-equality-around-the-world

relations that underlie it. Social exclusion overlaps with poverty but extends beyond it by explicitly embracing the relational and distributional aspects of poverty. It is crucial to understand that M&E practitioners work in accordance with their position within a society and its values, in an evidence landscape that is fixed (i.e. global indicators) but fluid (i.e. the domestication of global indicators).

Work within a monitoring environment can be experienced as a one-dimensional technocratic exercise, with it being a mechanistic and functional "tick box" process of managing data inflows and assessing progress against indicators in performance frameworks. Every piece of human data tells a story or frames a real, lived experience. In viewing ourselves as development M&E practitioners, we can interact with data in ways that detail lived realities in a world inundated by inequality on multiple levels. Following Shanker (2018), the equity-driven M&E practitioner can reflect on and pose challenging questions about data, such as how a data point categorising a person as "unemployed" is informed by how that individual is racialised, classed, gendered, and sexualised. Having the necessary granularity of evidence to do these kinds of equity-informed analyses becomes critical in driving a more ethical and developmental practice.

With some issues that inform an equity-oriented M&E approach being outlined previously, the following consideration focuses on the means by which these approaches can be translated into reimagined practices. Efforts to introduce equity as a key monitoring component are relatively recent and have yet to be systematically mainstreamed into government monitoring and evaluation processes. Much of the innovation and experimentation in this regard is being tested in the development and philanthropic sectors, and to some extent in the private sector. While the theoretical underpinnings of intersectionality provide an essential lens for understanding social exclusion challenges, the M&E sector needs to grapple with the "how to do equity" conundrum. The focus within an African context lies in the move to "decolonise" evaluation, which entails a process of comprehensive review of the ideological and epistemological underpinnings of current practices and knowledge. These should be gathered, taken apart, broken up, and critically examined to identify their benefits and weaknesses in terms of equity-informed practice. This is an emerging area of interest within the M&E field, and it represents opportunities for technical experimentation and innovation. Bowleg (2008), for example, argues that a key dilemma for researchers and M&E practitioners working with intersectionality is to grapple with the assumptions inherent in measurement as well as qualitative and quantitative data analyses, and to transform them so that equity can be integrated into practice.

In the M&E environment, an inherent bias exists towards easily quantifiable, pragmatic, and reductionist data that simplifies the process of making sense of a complex world. However, this approach contradicts the central tenet of intersectionality, which is that social identities and inequality are interdependent for many groups of people and are not mutually exclusive. Programme interventions and their associated M&E frameworks rely on output and outcome level indicators to measure progress and do not attempt the building of intersectionality thinking into the process. Discrete and continuous variables, such as gender, race, ethnicity, and age, are measured as independent factors but attempts to understand how these variables intersect and interact with each other are rarely undertaken. Considering this gap, interpretation becomes one of the most substantial tools in the methodological toolbox of researchers, data analysts, and those working with M&E systems. Significant progress is being made, for instance, in the United Kingdom, where charters have been put in place that require government institutions across the board to collect equity and inclusion data that ethically reflects intersectionality considerations.

Equity-oriented practitioners across the policy, programming, monitoring, analysis, and reporting spectrums can be sensitised and informed by a range of approaches such as heuristics, assessment tools, checklists, and matrices. Second and third-wave feminists and queer theorists working in gender and sexuality studies carried out much of the initial analytical thinking regarding embedded socioeconomic and political inequity. One such theorist, Caroline Moser, introduced the Gender Analysis Framework in the 1980s as a means of supporting an integrated gender-planning perspective in all development work, concentrating on the power relations between men and women. This approach allowed planners to engage with the complexity of inequality and introduced the subordination of women into planning and monitoring discourse.[10] Some of this analysis has been translated into public and development theory and practice. The gender-responsive assessment scale, a tool developed by gender studies theorists, enabled policymakers, programme designers, and M&E practitioners to reflect on and assess the level of gender-responsiveness manifested policies, strategies, and programmes. This also resulted in the interrogation of the kinds of indicators being used to collect critical data. The value of these analytical tools lies in their ability to be adapted to the assessment of gender-responsiveness as well as other areas of inequality, complimented by an additional intersectionality lens. The following scale is adapted from the World Health Organisation (WHO) Gender Responsive Assessment Scale, which was developed for managers in the health sector:[11]

---

10   European Institute for Gender Equality. 2019 available at https://eige.europa.eu/sites/default/files/mh0319271enn_002.pdf

11   https://www.who.int/gender/mainstreaming/GMH_Participant_GenderAssessmentScale.pdf

Table 2.1:   *WHO, Gender Responsive Assessment Scale*

Equity Responsive Assessment level	Consequences for policies, strategies, and programmes	The extent to which an Intersectionality Lens is applied
Equity-Unequal	• Perpetuates inequality by reinforcing unbalanced norms, roles, and relations. • Privileges one or more groups over others. • Often leads to one group enjoying more rights or opportunities than the other.	• There is no understanding of differences within groups or how social norms determine these differences. • Groups are homogenised and essentialised.
Equity-Blind	• Ignores societal norms, roles, and relations (gender, religion, culture, etc.). • Very often reinforces identity-based discrimination. • Ignores differences in opportunities and resource allocation for different groups/ populations. • Often constructed based on the principle of being "fair" by treating everyone the same.	• Poorly informed understanding of different identities within groups is not accounted for. • Intersections of identity and experiences are not seen. • Different identities and experiences are collapsed into generalised categories.
Equity-Sensitive	• Considers societal norms, roles, and relations. • Does not address inequality generated by unequal norms, roles, or relations. • Indicates equity awareness, although often no remedial action is developed.	• Some understanding of how social norms construct identities. • Some understanding of how intersecting identities determine people's life experiences. • Limited understanding of how to respond to intersectionality at a policy or programme level.

Equity Responsive Assessment level	Consequences for policies, strategies, and programmes	The extent to which an Intersectionality Lens is applied
Equity-Specific	• Considers societal norms, roles, and relations for different groups and how they affect access to and control over resources. • Considers the specific needs of different groups. • Intentionally targets and benefits a specific group of people to achieve certain policy or programme goals or meet certain needs. • Makes it easier for people to fulfil duties that are ascribed to them based on their perceived socioeconomic roles.	• A more nuanced understanding of how intersecting identities determine access to and control over resources. • Consideration is given to complex identity formation within groups. • Some ability to respond to intersectionality in the formulation of policies and programmes.
Equity-Transformative	• Considers gender norms, roles, and relations for different groups and that these affect access to and control over resources. • Considers the specific needs of different groups. • Addresses the causes of inequities. • Includes ways to transform harmful social norms, roles, and relations. • The objective is often to promote equality. • Includes strategies to foster progressive changes in power relationships between different groups of people.	• A strong theoretical grasp of intersecting identities and differential outcomes within and between groups. • Policy/programme responses consider intersectional impacts. • Addresses the causes of inequities at a granular level. • Includes ways to transform the modalities of intersectionality that negatively impact people's life experiences and outcomes.

An assessment scale such as this needs to be utilised across the policy and programme cycle to be effective. A challenge often faced by equity-responsive monitoring practitioners is the immensely complicated means of measuring the performance of strategies and programmes which have failed to integrate equity and intersectionality considerations during the conceptualisation and design phase. However, monitoring systems can be proactive in inserting equity adaptations from the bottom up. A recent application is in the public health field of HIV/AIDS and

human rights, where the Global Fund (2019) has, in its "Breaking Down Barriers" initiative, set out programming parameters to address inequities in access to critical health and legal services.[12] In many countries in Southern Africa, HIV has become a manageable health condition. However, prevalence remains high in vulnerable populations, including the LGBTQIA+ community, sex workers, drug users, migrant communities, as well as vulnerable girls and women, where stigma and discrimination can impede access to public health and legal services. This example illustrates how the consideration of modifiable societal and contextual factors in the reduction of health inequalities becomes increasingly important both to implement and measure.

From the data generation and monitoring perspective, the challenge is that many aspects of individual lives remain largely unaccounted for in public data sources such as national censuses, household surveys, and demographic and health surveys. In the case of both communicable diseases (HIV, TB, and COVID-19) and non-communicable diseases (obesity, alcohol use disorder, and mental health) there are many reasons for this disparity which lie outside of the control of M&E practitioners. Some of these include the criminalisation of same-sex relations, sex work, drug use, migrancy, and their associated levels of societal stigma, discrimination, and marginalisation. Without reliable disaggregated data that reflects the intersectional experiences of people's lives, the challenge for policymakers and programmers lies in the development of targeted interventions that are appropriately evidence-informed and are designed to reduce or eliminate inequities in access and opportunity.

Moving to More Equitable, Intersectional, and Inclusive Measurement

A challenge faced by monitoring systems includes the collection of data from populations that are difficult to reach. These populations often view the government as a threat to their safety and well-being, making this problem particularly relevant in terms of the government data collection processes. Civil society organisations that specialise in working with vulnerable populations are taking innovative strides in data collection from these populations. One such organisation, Frontline AIDS, has developed a peer-driven data collection system called REAct which monitors human rights violations and focuses on HIV programming as well as advocacy for marginalised people who are vulnerable to or affected by HIV and AIDS.[13] Academic research institutions may also undertake Integrated Biological and Behavioural

---

12   The Global Fund. Human Rights https://www.theglobalfund.org/en/human-rights/#:~:text=Through%20our%20Breaking%20Down%20Barriers,HIV%2C%20TB%20and%20malaria%20services.

13   The technicalities of the REAct system are set out in the REAct User Guide https://frontlineaids.org/resources/react-user-guide/

Surveys (IBBS) that use peer researchers to collect personal data from hard-to-reach populations such as at-risk ethnic minorities, transgender people, sex workers, and survivors of rape. These approaches effectively use bespoke data collection tools to collect data that accurately reflects the intersectional reality of individuals who report into the system and makes referrals and tailor-made responses easier. The REAct system, for example, is run by peer community health workers and paralegals accepted and trusted by vulnerable people. They can collect data that official data collection methods cannot and feed it into national data collection systems such as health information systems. These applications in health research mark an emerging research paradigm that seeks to move beyond single or typically favoured categories of analysis (i.e., sex, gender, race, and class) to consider simultaneous interactions between different aspects of social identity, as well as the impact of systems and processes of oppression and domination (Palmen 2021).

The intersectionality wheel shown in Figure 3 (WHO 2020) is a useful heuristic that enables practitioners to think about what intersectionality means in practice. It uses a variant of the socioecological model that is premised on the overlapping and intersecting domains of the individual, the community, the society, and the structural dimensions of that society.

Figure 2.3:   *World Health Organisation, Intersectionality Wheel, s.a. (World Health Organisation 2020)*

The intersectionality wheel illustrates how multiple individual characteristics (such as age, gender, and education) interact within broader processes of social discrimination (such as ableism, sexism, and racism) and structural barriers (such as political, class, and economic) to shape an individual's position within a given society. It allows researchers to model demographic characteristics across different levels by capturing people's lived experiences within the intersecting macro, meso, and microsystems. Intersectional approaches seek to consider the positions of all members of a given society and aim to illuminate the position of the most marginalised within existing power dynamics through the lens of the barriers faced in accessing equitable public services.

Government agencies in certain parts of the world actively use intersectionality as a monitoring strategy by addressing identity stereotypes and individuals' unique and intersecting experiences that result from race, gender, sexuality, disability, or other forms of vulnerability. Women and Gender Equality Canada has developed a process called Gender-based Analysis Plus (GBA+), which is an analytical process that can be used to analyse the "gendered" aspects of Canadian government policy to assess the different experiences of women, men, and non-binary people within policies, programmes, and initiatives.[14] The 'Plus' component considers the intersectionality of identity factors and how the relationships between these identity factors impact the way in which government programmes and initiatives are designed, monitored, and experienced. The Ontario Human Rights Commission in Canada argues that intersectional analysis should become one of the lenses through which the social context of the individual can be examined and, in some measure, should assist in addressing social conditions relating to poverty, low income, and homelessness through improved data (OHRC 2021). In countries such as Canada, the United States, New Zealand, Australia, Mexico, and South Africa, First Nation people are often the most marginalised and stigmatised population segments.

Developed in the United Kingdom in 2005, the Athena Swan Charter is a globally used framework used to support and transform gender equality within higher education (HE) and research. Initially established to encourage and recognise a commitment to the advancement of the careers of women in science, technology, engineering, mathematics, and medicine (STEMM) employment, it is now being used in many countries to address gender and identity equality more broadly, no longer focusing solely on progression that affects women (AdvanceHE 2021). The South African Workplace Equality Index (SAWEI) is a similar monitoring system that enables South African companies to measure their progress against independently determined and research-based best practices regarding LGBTQIA+ inclusion in the workplace. This system uses a survey consisting of six elements that measure equity in the structures

---

14   Government of Canada. Women and Gender Equality Canada available at https://women-gender-equality.canada.ca/en.html

and behaviours of companies.[15] The Gender Lens Investing Initiative is an intervention that stems from the Global Impact Investing Network (GIIN) which supports private sector impact investors in the active integration of a gender lens strategy into their investment portfolios to address the systemic issue of a lack of gender equity that has been present in the business and investment community throughout its existence.[16] Part of this initiative has been to set up monitoring systems that can track the impact of investments on the lives of women and girls. The International Foundation for Electoral Systems (IFES) has developed a new assessment framework to identify intersectional barriers and opportunities related to people's political participation with multiple social identities, including gender, disability, and age.[17] This participatory tool was developed based on the recognition that marginalised populations, including women, persons with disabilities, LGBTQIA+ persons, and people from ethnic or religious minorities, experience systemic discrimination in many countries. In some cases, victims of physical or psychological violence have had their political and constitutional rights undermined by voting rights barriers (Atkinson 2018).

Although marginalised groups experience well-documented barriers to political participation, there remains little analysis of the intersecting challenges or common goals between individuals in each area. The M&E field has recently undergone growth through the emergence of The Equitable Evaluation Initiative (EEI), which has assisted in the advancement of equity and the expansion of rigorous notions of validity and has fuelled the embracing of complexity.[18] This approach shifts the current evaluation paradigm and encourages mindsets and practices to evolve towards a more equity-oriented practice. Its Equitable Evaluation Framework supports M&E practitioners working with Foundations and Civil Society Organisations to replace conventional narratives that marginalise, minimise, and disrespect people with ones that reflect an understanding of the systemic and structural barriers that limit the possibilities of people and their ability to thrive. These methodologies provide equity-focused tools for data collection that allow people who identify with multiple social identities to share the different ways in which they participate civically and politically, prioritise identified solutions to barriers, and share their opinions in targeted surveys. Where used purposely, these methodologies are a critical first step

---

15   The SAWEI is managed by the LGBT+ Management Forum that works with companies across South Africa to create safe and equitable workplaces that enable lesbian, gay, bisexual, trans (LGBT+) professionals to contribute to their fullest potential, and the index reports are available at http://lgbtforum.org/news/view/sawei-2021-results-announced-with-5-gold-tiered-companies

16   Global Impact Investing Network,. Gender Lens Investing Initiative https://thegiin.org/gender-lens-investing-initiative

17   International Foundation for Electoral Systems. Participation and Inclusion https://www.ifes.org/issues/participation-and-inclusion

18   Equitable Evaluation Initiative https://www.equitableeval.org/

towards generating intersectional data that can feed into and augment other more standardised datasets. UNESCO's World Inequality Database on Education (WIDE) is another useful interactive tool for illustrating how different forms of disparity such as gender, household wealth, ethnicity, religion, and residence, play an important part in shaping opportunities for and access to education.[19] It visually demonstrates how overlapping disadvantages can compound education disparities. The World Inequality Database was initially created as the The World Top Incomes Database (WTID) in January 2011 with the aim of providing free and convenient access to all the existing series. It then expanded to include a series on income inequality for more than 30 countries which spanned over most of the 20th and early 21st centuries, with over 40 additional countries now under study.

Applying an intersectional approach (or lens) helps assess the potential impacts of initiatives – positive or negative - based on their multiple identity factors, enabling programmers to identify risks and potential challenges early and create mitigation strategies in response. An intersectional approach should be applied at all stages of an initiative, from development through implementation and monitoring to evaluation. The consideration and identification of people's diversity and multiple identity factors assist practitioners with the innovation and consideration of issues and policies in different, more reflective ways. This has clear implications for the design of programme monitoring systems. Most M&E systems, whether in government, the development sector, or civil society, are generally set up for administrative compliance, accountability, and routine reporting purposes, and tend to focus on broad policy and programme outputs that are easily measurable and reportable. This is crucial for government accountability to the type, scale, cost, and reach of their services to citizens. However, in the developing world context, most public services must address significant socioeconomic disparities at a broad societal level, such as addressing poverty, unemployment, and inequality. In most instances, their monitoring systems are not optimally equipped for tracking progress on gender and social inclusion and cannot, as a result, provide policymakers with data that is sufficiently granular and disaggregated. This could be a capacity issue, a design issue, a practice issue, or a combination of all three. In some cases, there may also be a lack of political appetite for collecting data on certain population segments.

For an effective monitoring system to be responsive to gender and social inclusion, the capacity, design, and practice issues need to be reappraised and reengineered. A more responsive monitoring system should provide more refined data to highlight the equity gaps and enable managers to better assess progress on key equity, gender, and social inclusion concerns, leading to more evidence-informed policy and programme decisions. There is also a need to have a more focused emphasis on advocacy and mainstreaming work with practitioners if there is a vision that

---

19   UNESCO https://learningportal.iiep.unesco.org/en/blog/one-way-of-measuring-marginalization-in-education

they can become developmental monitoring and evaluation specialists rather than M&E technicians. In reality, sourcing and producing standard data is not enough to make that data useful in terms of addressing equity considerations. Citizens are not interested in data; they are interested in improved services being designed and delivered based on available data and information that surfaces their realities. If that data is inadequate and only partially inclusive of a full range of citizen experiences, then the evidence bases for building, strengthening, and sustaining equitable services are compromised. Conventional monitoring systems within the government draw on available data ecosystems and produce compliance data or data that informs the indicators used for performance planning. The challenge lies in the data as well as M&E practitioners themselves being "blind" to the equity gaps in the data and the inadequacy of the data in informing an understanding of intersectionality issues. Efforts to address these equity gaps across the government policy and planning cycles are manifested in initiatives such as gender-responsive budgeting. This is a growing area of interest for developing country governments that are aware of the need to address gender inequalities across all sectors. In South Africa, for example, the Gender Responsive Planning, Budgeting, Monitoring, Evaluation, and Auditing Framework (GRPBMEA) is being championed by the National Treasury.

In some cases, grounds such as sex, race, ethnicity, or disability may intersect and produce unique effects creating "discrete and insular minorities" who are socially handicapped due to these same characteristics (OHCR 2016). Simple examples here include the situation of undocumented young migrant women, sex workers, or transgender women whom the state may deny critical sexual and reproductive health services due to their liminal status. Such situations may be typically thought of as people "falling through the cracks" or "being left behind" by government services that are intended to be inclusive and equitable. The lack of sufficiently disaggregated data (or no data at all) on these populations highlights the importance of monitoring systems of creating mechanisms that reflect and capture individual and group experiences based on multiple identities linked to more than one ground of discrimination. Alternatively, any one of these characteristics may intersect with other grounds such as social assistance, family status, and a further link to economic and social and class status to create unique experiences for the individuals that current developmental frameworks ignore. Even when combined with other grounds such as social assistance and family status, the extent of the discrimination may not be revealed by a traditional, non-intersectional approach.[20]

An intersectional analysis can be informed by developments in gender equality analysis, critical race analysis, disability rights analysis, and equality rights juris-prudence. These strategies have been developed to address the stereotypes and

---

20   Ontario Human Rights Commission https://www.ohrc.on.ca/en/intersectional-approach-discrimination-addressing-multiple-grounds-human-rights-claims/applying-intersectional-approach

unique and intersecting experiences of individuals because of race, gender, or disability. They would form a necessary part of the contextual and analytical framework. However, Runyan (2018) warns that a tokenistic approach could obscure the transformative intention of this approach and dilute the nature of inequalities without the impetus to use monitoring data to address the underlying structures that produce and sustain injustice. The relevance of this for monitoring systems is that a lack of critical equity data should be a concern for practitioners and management staff, as programmes with weak monitoring systems are not likely to be easily evaluable in terms of impact. Referring specifically to international development work Riddell (2014) argues that there is a lack of reliable, consistent, and robust programme information to answer whether development interventions work, partly because of inadequate baseline data and the often-weak monitoring upon which the assessment of impact is based.

A case in point is the stubborn persistence of gender-based violence (GBV), which is an increasingly visible societal challenge but is also one that yields concerningly little data. Accurate and reliable generalised data on the prevalence and incidence of GBV has been notoriously difficult to collect. Therefore, it is challenging to analyse available data for evidence-driven and equity-oriented programming. The intersectionality lens disrupts this problem further by calling for a greater understanding of how GBV impacts women and girls based not just on their gender but also on intersecting identities such as race, class, locality, and sexual identity. Based on this, much work is needed to make monitoring systems more inclusive and more responsive to equity considerations. A key consideration arising from the GBV challenge is that data "talks about" victims and survivors rather than "talks with" them. The means and extent to which practitioners decide to develop monitoring systems, including indicators, depend on their specific aims. They can simply collect data, such as police statistics, about a particularly concerning societal issue such as GBV, or they can go a step further by making sure that primary stakeholders, particularly those who are typically marginalised or excluded, participate in some way in to inform monitoring processes and practices so that their experiences and understandings feed directly into social transformation agendas.

Building an Equity-Oriented M&E Practice

This chapter has attempted to outline some of the theoretical underpinnings of an equity-oriented M&E practice and has pointed to some of the initiatives being undertaken worldwide to pilot innovative approaches. For those training, working with, or mentoring emerging M&E practitioners, several critical interventions can lead to an impact in strengthened interest in and commitment to the theory and praxis of equity-led monitoring systems.

Building a Commitment to Equity

As technical systems, monitoring systems will be as effective or ineffective as allowed by their design and the ability of their practitioners. Ideally, the design of monitoring systems should be informed by equity issues. Monitoring systems also need to be appropriately geared to tracking the routine outcomes of policy and strategy provisions on equity issues of national concern at both the macro and the micro level. Examples of these include gender equality issues related to remuneration, financial independence, political participation, safety and security, sexual and reproductive health, and bodily integrity. Programme designers and M&E specialists are generally well-attuned to equity issues in the development sector, as their work often focuses on vulnerable, marginalised, and at-risk populations. Even though government service delivery in developing countries carries an immense equity burden, public sector agencies often lack adequate sensitisation and awareness around effective integration of equity and intersectionality considerations into their work. Monitoring systems may be characterised by data gaps that result in the needs of specific population segments being unintentionally overlooked or intentionally ignored. As an example, many countries see ethnic and religious minorities, political activists, members of the LGBTQIA+ community, and sex workers being heavily stigmatised, marginalised, or even criminalised, and deliberately excluded from national statistics. Emerging practitioners must learn to understand how power dynamics include and exclude people and how this impacts the quality and availability of equity-informed data.

Awareness-Raising, Values Clarification, and Responsibilities as a Public Servant

Public services operate as significant sources of employment within countries and have extremely diverse workforces. The ideal scenario from a public service and administration perspective is to ensure that those working as public servants are inducted, capacitated, and assessed according to conditions of employment and prevailing policies and regulations. In the South African context, this requires that public servants at all levels comply with constitutional mandates on core human rights and equity issues and that equity is integrated into planning and programming initiatives.

Government in-service capacity development programmes on equity-related issues are well established, although their efficacy has not been thoroughly evaluated. In South Africa, the National School of Government runs a range of in-service training programmes for public servants, including ones that cover issues such as human rights and gender mainstreaming, but the impact of many of these programmes is yet to be evaluated. However, it has created a context for organically building on work that has previously been conducted within public services to mainstream

gender as standard practice across all sectors and disciplines of government. For M&E practitioners, this should arguably have resulted in a better understanding of how to work with sex-disaggregated data and undertake gender-responsive data analysis work. If this assumption is correct, these understandings can be extended to include an equity and inclusion practice that uses intersectionality as a core monitoring tool. This involves the sensitisation of M&E practitioners regarding the value of increasing equity dimensions in their data products by using intersectional analysis in a strategic and targeted way.

The result of this process brings about an additional challenge that sees individuals within the workforce lacking the knowledge, skills, and integrity to proactively address certain equity issues that form a part of their work mandate. An example of this involves cases of healthcare workers who possess personal value systems, life experiences, and cultural backgrounds which contradict their obligation to provide non-discriminatory and non-stigmatising services. In many countries, Ministries of Health provide values clarification training for healthcare workers to sensitise them to human rights issues, inclusive care, and equitable health service provision. These kinds of interventions give healthcare providers a chance to examine their values, attitudes, and beliefs and change attitudes that hinder the service progress, while honing professionalism among the health workers. Perhaps the most challenging example here is the constitutional right that healthcare providers have in countries such as the United States and South Africa to refuse to provide legally-sanctioned abortions based on freedom of conscience, religion, and opinion. The same concerns could apply to M&E practitioners who may already have a negative predisposition to collecting and analysing monitoring data relating to specific population groups or actions. Working with equity and intersectionality issues presupposes that the following preconditions should ideally be in place:

- Pre-service and in-service education and training to provide adequate knowledge and understanding of equity issues integral to effective and efficient monitoring systems.

- Strengthening a theoretical understanding of how data can be used to drive equity-informed monitoring systems.

- Values clarification training is provided as a standard practice to support staff in reconciling personal value systems with the requirements of their job responsibilities.

- M&E managers are sufficiently capacitated to provide leadership, guidance, and mentorship to staff in applying equity-focused monitoring systems and using intersectionality and inclusion frameworks.

Building Technical Expertise

Typically, monitoring systems work with linear and unidimensional data points, which do not necessarily provide insight into the complex ways that inequalities operate at the individual, societal, and structural levels. Therefore, focusing on identities is only a part of the work that an equity-oriented practice requires. A composite indicator, such as the one used to populate the UN's Human Development Index, ranks countries annually based on their level of human development and is informative but incomplete as it reflects the effects of inequalities rather than the causes thereof.[21] More difficult questions need to be addressed around structural inequality and the dynamics that create and maintain systems of identity-based inequality. This means understanding the drivers of inequality and the consequences of inequity in the real world for specific populations of people and how data does or does not evidence this. In the context of monitoring systems, practitioners need specific guidance and examples about how to import inclusion and intersectionality as analytical frameworks into their practice. This shift requires an express connection of individual experiences of discrimination with the structures of power and exclusion from which discrimination is bred. Reflexive consideration is required by practitioners of the systems and processes that they use, and how these may, by virtue of their design, perpetuate the use and application of equity-blind data.

An equity-informed framing of the monitoring practice can lead to a more analytical M&E practice that recognises groups of people as having multiple diverse and intersecting identities that impact how they understand and experience government initiatives and services. Public sector monitoring functions are normative and rule-based, providing routine evidence sets that feed into programme management processes. This iterative data collection process informs government decision-making in the context of policy-making programming and implementation. Data generated through statistical and M&E processes will routinely be disaggregated along a limited spectrum of variables, among other things gender, age, race, location, and employment status. A monitoring and evaluation officer coming into this environment will understand their role within the framework of such a normative, rule-based system and may not necessarily be aware of the wider potential of monitoring systems for driving an equity-informed practice.

Building technical capabilities to operationalise equity-informed practices should be integrated into pre-service and in-service education and training. This needs to happen in a very practical way, using approaches that have been designed in a range of different sectors to address equity considerations. Generally, researchers

---

21   The Human Development Index (HDI) is a summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable and have a decent standard of living. The HDI is the geometric mean of normalized indices for each of the three dimensions. bit.ly/3rVbwTT

have relied primarily on qualitative methods to investigate intersectionality, raising questions about the epistemological and methodological possibilities of generating and using quantitative intersectionality data. In this regard, practitioners could engage with studies such as the one published in the New England Journal of Medicine that explores how gender and race influenced the referral of patients for cardiac catheterisation (Shulman et al. 1999). They had data on both race and gender and produced a statistical analysis. They did a "main effects" analysis in which they looked at the influence of gender, followed by looking at the influence of race. They then combined these main effects additively and created a chart that looked like the one in Figure 4 seen below:

Figure 2.4: *Bowleg and Bauer, Using a quantitative intersectionality-based model, 2016*

Bowleg and Bauer (2016) have used this example to point out that adding together main effects such as gender and race does not produce an intersectional analysis. A different set of results is found when the analysis is redone using a quantitative intersectionality-based model. It quickly becomes apparent that the actual bias here is against black women. The initial incorrect results showing the lower odds ratios for white women and black men in the additive model included bias being shown to black women. As Rouhani (2014) notes, this happens because quantitative researchers traditionally seek to address issues of social inequity by investigating axes of inequity, such as race, gender, class, and sexuality, and only considering the potential interconnectedness of these axes. To build an intersectional model,

practitioners need to move beyond an additive model or algorithm that is built, for example, as **Outcome = Race + Gender + Sexual Orientation**. This model can be used to understand the effects of one of the predictors, such as sexual orientation, on the outcome while holding other predictors constant. Holding gender constant while looking at the impact of race and sexual orientation does not, however, tell us whether the impact of these individual characteristics differ when they are allowed to fluctuate. Alternatively stated, what is the effect of sexual orientation when gender is allowed to be either male or female? This is the nature of core questions of intersectional analysis. Intersectionality posits that experiences at an intersection are co-constituted and must be considered jointly. This distinction between additive and intersectional approaches maps onto quantitative distinctions between main effects and heterogeneity of effect (Hancock 2007). By adding an interaction term (essentially, multiplication), the new model now looks like **Outcome = Race * Gender * Sexual Orientation**. This model can answer how the outcome changes for different combinations of the variables by looking at all three predictors together rather than individually. The multiplication in the model accurately estimates the simultaneous and layered effects of the different variables. In an intersectionality-informed analysis, the additive approach is incorporated as an initial 'baseline' model, upon which further analyses are applied using multiplicativity to account for the conditional effects of intersecting categories on a social outcome (Rouhani 2014).

A key principle of equity in data and intersectional analysis involves acknowledgment that the inclusion of individual-level data alone frequently produces biased and incorrect results. Without any context, data analysis can yield incorrect and skewed data outcomes. For example, we may have national or sub-national level demographic data that tells us about the number of unemployed young women in the 18-to-24-year cohort, disaggregated by race. However, it tells us little about these women beyond those core variables. The building of an intersectional model involves taking additional steps to look at variables and data that measure the context and communities in which individuals are situated. For instance, in a model about the effects of age, gender, and refugee status on educational outcomes, it is important to include measures of how accepting each community is to refugees. It could also include variables measuring the availability of education in various languages, systemic regulations on gender and education, local levels of xenophobia, and other potential forms of human rights violations. Various statistical methods enable this type of analysis, the most common of which are multilevel models designed to include variables measured at the individual level and several broader levels of aggregation, such as community and country (Sage 2020). Like traditional regression, these models can include multiplication, not just addition, as posited by the intersectionality paradigm.

In her Primer on intersectional analysis, Rouhani (2014) includes an example exploring the intersectional effects of race, education, and urban area. The different urban areas have different policies, so this variable acts as a measure of structural level influence in the model. Researchers could conduct cross-contextual comparisons that would evaluate the impact of this policy across urban settings to empirically investigate how policy constructs the relative power and privileges within the system. This could be done through a comparison on cumulative years before and after the policy introduction throughout cities and states that enacted the legislation versus those that did not.

Table 2.2: *Rouhani, Applying an intersectional multiplicative approach, 2014*

		Odds Ratio			
		Chicago	Los Angeles	New York	San Francisco
Race X education interactions					
Black	OR less than high school	3.280	1.752	2.580	4.850
	OR high school graduate	1.896	1.200	1.950	2.540
	OR some college	1.461	1.390	1.428	2.102
	OR college graduate (ref)	1.000	1.000	1.000	1.000
White	OR less than high school	2.133	1.450	1.890	3.504
	OR high school graduate	1.621	1.325	1.450	2.320
	OR some college	1.181	1.120	1.320	1.950
	OR college graduate (ref)	1.000	1.000	1.000	1.000

There are packages and examples of multiplicative and multilevel models that can be used by M&E practitioners to add an intersectional analysis to their work which are available in most standard statistical packages. There is one example using R on multiplication or interactions as well as one on multilevel modelling. For SPSS users, there is one on interactions and one on multilevel modelling as well.

In reality, individuals are not single data points (e.g., male or female, literate or illiterate), and no individual lives one of their characteristics at a time. Our lived experiences are a simultaneous combination of all our characteristics and experiences. If we want to use data to reflect people's lived experiences more accurately, then we need to use it in combination with other data sets. This is a multiplicative approach and allows us to see how different characteristics are interacting or intersecting and then enables us to better analyse the positionality of an individual within a society. The use of multiplicative models allows practitioners

to see how different characteristics are interacting or intersecting (Krause 2021). When this frame is used as an equity-oriented monitoring tool, the data shows that cumulative marginalisation is much stronger than the marginalisation of one group at a time. When individuals are members of more than one marginalised group, the effect can be cumulative, as illustrated in Figure 5 below:

Figure 2.5:   *Krause, Applying the multiplicative approach to demonstrate cumulative marginalisation, 2021*

This kind of mapping demonstrates the effects of cumulative marginalisation. The more marginalised identities a person holds, the higher the order of intersection and the lower their probability of being treated equitably. Figure 5 looks at the cumulative marginality of women and the likelihood of their engagement in political activism. From the chart, it is possible to see that a woman has a probability of 17% regarding active involvement. The probability of a woman from the lowest social class getting involved is 9%. However, a woman from the lowest social class and belonging to the marginalised ethnic group has a 7% probability of getting involved. In this example, data is used multiplicatively instead of additively to highlight the lived experience of cumulative marginalisation. Using complex, multi-layered demographic data

in research and monitoring systems is not always simple. However, it is crucial to ensure that the work being done is equitable and that it gives practitioners a more granular understanding of the issues they are working with.

Conclusion

Arguably, M&E systems within government are slow to respond to trends and innovations around M&E within development, civil society, and academic sectors. The major players in the development field, including the World Bank, the regional banks, the OECD, the UN agencies, and international NGOs, are undoubtedly leaders in the development of equity-responsive monitoring and evaluation systems. In many cases, they work alongside government entities to strengthen monitoring systems in raising efficacy in the collection, processing, and utilisation of equity-informed data. Development agencies and other academic and civil society organisations working at the forefront of monitoring and evaluation innovation can only do so much. Governments have the mandate and the obligation to provide for their most vulnerable and marginalised populations through evidence informed and equity-driven policymaking and programming. Government M&E systems in Africa can also draw on cutting edge work that is being done through the auspices of organisations such as AfrEA and Twende Mbele on the development of indigenous and Afrocentric M&E practices that focus on equity as a core epistemological framework.

This chapter has attempted to provide reflections on how practitioners working in the public service can broaden their thinking about their practice and has suggested how routine monitoring systems can expand their remit by integrating methodologies for measuring equity, inclusion, and intersectionality.

Reference List

AdvanceHE. 2021. *The Athena Swan Charter* https://www.advance-he.ac.uk/equality-charters/athena-swan-charter

Agrawal, R., Kalugampitiya, A., Rinxin, J. and Hashim, K. 2017. Leveraging Monitoring and Evaluation Systems for Good Governance. *Evaluation for Agenda 2030: Providing Evidence on Progress and Sustainability.* Exeter, UK: IDEAS http://web.undp.org/evaluation/documents/Books/Evaluation_for_Agenda_2030.pdf

Atkinson, V. 2018. *New Intersectionality Assessment Framework. International Foundation for Electoral Systems* https://www.ifes.org/news/new-intersectionality-assessment-framework

Bok, K. 2018. Definitions and Measurement of Social Exclusion - a Conceptual and Methodological Review. *Advances in Applied Sociology,* Vol.8 No.5, 2018 https://www.scirp.org/journal/papercitationdetails.aspx?paperid=84529&JournalID=1002

Bowleg, L. 2008. *When Black + Lesbian + Woman ≠ Black Lesbian Woman: The Methodological Challenges of Qualitative and Quantitative Intersectionality Research. Sex Roles* 59, pp. 312–325 https://doi.org/10.1007/s11199-008-9400-z

Bowleg, L. and Bauer, G. 2016. Invited Reflection: Quantifying Intersectionality. *Psychology of Women Quarterly* 2016, Vol 40(3), pp. 337-341 https://journals.sagepub.com/doi/10.1177/0361684316654282 .

Corus, C., Saatcioglu, B., Kaufman-Scarborough, C. Blocker, C., Upadhyaya, S. and Appau, S. 2016. Transforming Poverty-Related Policy with Intersectionality in *Journal of Public Policy and Marketing.* vol 35(2) Fall 2016 211-222 https://www.jstor.org/stable/44164853. https://doi.org/10.1509/jppm.15.141

Crenshaw, K. 1989. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics in University of Chicago Legal Forum Volume 1989, Issue 1 Article 8. https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1052&context=uclf

Crenshaw, Kimberley. 1994. "Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color", in The Public Nature of Private Violence, M. Fineman and R. Mykitiuk Eds. Routledge: New York. pp. 93-118 http://www.hsph.harvard.edu/Organisations/healthnet/WoC/feminisms/crenshaw.html

Fischer, A. 2011. *Reconceiving social exclusion.* Brooks World Poverty Institute, The University of Manchester http://hummedia.manchester.ac.uk/institutes/gdi/publications/workingpapers/bwpi/bwpi-wp-14611.pdf

Frontline AIDS. 2019. REAct User Guide https://frontlineaids.org/were-on-the-frontline-of-human-rights/

Hancock, A-M. 2007. When multiplication doesn't equal quick addition: Examining intersectionality as a research paradigm. Perspectives on Politics, 5. bit.ly/45gxH5l. https://doi.org/10.1017/S1537592707070065

Khan, S., Combaz, E. & McAslan Fraser, E. 2015 *Social exclusion: topic guide.* Revised edition. Birmingham, UK: GSDRC, University of Birmingham https://gsdrc.org/wp-content/uploads/2015/08/SocialExclusion.pdf

Krause, H. 2021. *Why We Need Intersectionality in our Demographic Data. Current Events, Data Analysis Concepts.* https://idatassist.com/why-we-need-intersectionality-in-our-demographic-data/

Lorgelly, P., Lorimer, K., Fenwick, E., Briggs, A. 2008. *The Capability Approach: developing an instrument for evaluating public health interventions.* University of Glasgow Section of Public Health and Health Policy https://www.gcph.co.uk/assets/0000/0430/Capabilities_-_full_report__August_08.pdf

Mackay, K. 2007. *How to Build M&E Systems to Support Better Government.* World bank Independent Evaluation Group: Washington. https://www.samea.org.za/documents/Seminar%20documents/How_to_build_ME_gov.pdf. https://doi.org/10.1596/978-0-8213-7191-6

Mbele, T. 2021. *What is meant by transforming evaluations for Africa?* https://twendembele.org/wp-content/uploads/2021/04/MAE-Practice-Brief-Final.pdf

Mena, E., Bolte, G. 2019. The ADVANCE GENDER Study Group. *Intersectionality-based quantitative health research and sex/gender sensitivity: a scoping review.* Int J Equity Health 18, p. 199. https://doi.org/10.1186/s12939-019-1098-8

Morkel, C. 2021. *Made in Africa Evaluation: What is meant by transforming evaluation?* Working Paper. Johannesburg: Twende Mbele.

OECD. 2018. *Development Co-operation Report 2018: Joining Forces to Leave No One Behind*, OECD Publishing: Paris, https://doi.org/10.1787/dcr-2018-en.

OECD and SIGI 2021. *Regional Report for Africa, Social Institutions and Gender Index.* OECD Publishing: Paris, https://doi.org/10.1787/a6d95d90-en.

Ontario Human Rights Commission. 1 July, 2023. *Applying an intersectional approach* http://www.ohrc.on.ca/en/intersectional-approach-discrimination-addressing-multiple-grounds-human-rights-claims/applying-intersectional-approach

Palmen, R. 2021. *Act on Intersectionality.* Zenodo. http://doi.org/10.5281/zenodo.5602859

Riddell, RC. 2014. *Does foreign aid really work? An updated assessment.* Development Policy Centre Discussion Paper 33, Crawford School of Public Policy, The Australian National University, Canberra. https://doi.org/10.2139/ssrn.2409847

Rohwerder, B. 2016. *Poverty and Inequality: Topic guide.* Birmingham, UK: GSDRC, University of Birmingham.

Rouhani, S. 2014. I*ntersectionality-informed Quantitative Research: A Primer.* Institute for Intersectionality Research and Policy.

Runyan, A. 2018. What Is Intersectionality and Why Is It Important? Building solidarity in the fight for social justice. *American Association of University Professors* November-December 2018. https://www.aaup.org/article/what-intersectionality-and-why-it-important#.Yb9PdmhBzrc

Shanker, V. 2018. *What is intersectionality and how can evaluators do it?* American Evaluation Association. https://aea365.org/blog/fie-tig-week-what-is-intersectionality-and-how-can-evaluators-do-it-part-1-situation-analysis-by-vidhya-shanker/

Shulman, K. Harless, J. Kerner, J. Sistrunk, S., Gersh, B., Dube, R., Taleghani, C., Burke, J., Williams, S., Eisenberg, J., Escarce., J. 1999. The Effect of Race and Sex on Physicians' Recommendations for Cardiac Catheterization. *The New England Journal of Medicine.* 340, pp. 618-626. https://www.nejm.org/doi/full/10.1056/nejm199902253400806

Statistics South Africa. 2014. *The South African MPI: creating a multidimensional poverty index using census data.* Statistics South Africa: Pretoria. http://www.statssa.gov.za/publications/Report-03-10-08/Report-03-10-082014.pdf

Symington, A. 2004. *Intersectionality: A Tool for Gender and Economic Justice, Facts and Issues.* The Association for Women's Rights in Development (AWID).

The Global Fund. 2019. *Advisory Review Removing human rights-related barriers: Operationalizing the human rights aspects of Global Fund Strategic Objective 3*. OIG. Geneva. https://www.theglobalfund.org/media/9020/oig_gf-oig-19-023_report_en.pdf

The Opportunity Agenda. 2017. *Ten Tips for Putting Intersectionality into Practice.* https://www.opportunityagenda.org/explore/resources-publications/ten-tips-putting-intersectionality-practice

Chapter 2

Chapter 2: Strengthening Equity-Informed Monitoring Systems Within the Public Service

Philip Browne

Introduction

Governments worldwide implement results-based performance monitoring frameworks to measure and evaluate the programme progress and effectiveness, aiming to enhance performance and achieve strategic development priorities. These frameworks serve two interconnected and mutually reinforcing functions of good governance. Firstly, they ensure accountability by assessing state performance in relation to expenditure, guaranteeing that government services, funded by taxpayers, are accessible and equitable to all citizens. The routine M&E systems employed by governments across sectors tend to prioritise the compliance and accountability aspects of good governance, aligning with positivist requirements.

However, this chapter asserts that the other facet of good governance, namely the imperative to comprehend why significant portions of society remain underserved and marginalised, demands greater attention. The summary presented in Figure 1 illustrates that equity and inclusion must be integral characteristics of good governance (Agrawal Kalugampitiya, Rinxin & Hashim 2017).

Figure 2.1:   *Agrawal et al, The Place of Equity and Inclusion in Good Governance Systems. 2017*

In the context of public sectors, good governance encompasses the structures, systems, and government entities that collaborate to facilitate evidence-informed decision-making processes and the execution of policies, programmes, and projects aimed at fostering equitable economic and social development, ensuring that no one is excluded. At the core of public sector good governance lies the monitoring and evaluation process, which serves as a mechanism for performance management and accountability.

Within this expansive domain of good governance, there exist opportunities for leaders and practitioners in monitoring and evaluation to critically examine their practices and identify approaches to enhance monitoring systems with a focus on equity.

The first step involves mapping out the theoretical underpinnings of equity-informed monitoring, evaluation, and learning practice and explaining why this approach

goes beyond compliance to serve a more developmental purpose. This chapter will unpack some of these theoretical positions on the role of equity in M&E and examine how these understandings are being translated into emerging technical practices in different sectors. It will also address the types of capacities that need to be nurtured to achieve equity-oriented monitoring.

The public sector should question whether evaluation, as a practice, contributes to entrenching and maintaining conditions of inequality or if it holds the potential to act as a transformative developmental agent through equity-driven evaluation methodologies. Morkel (2021) argues that evaluation for transformational change or transformative evaluation should be grounded in social justice and equity in Africa. In her blog post, Jara Dean-Coffey, a specialist and writer in the field of evaluation, emphasises that "our distance from the work/people/issue/community has often led to us not seeing/feeling what is happening, but also not understanding the nuance and complexity that exists in the human experience, let alone the planet or universe".[2] The African Evaluation Association (AfrEA) acknowledges this omission and urges evaluators to consider that "evaluation considers issues and norms that are sensitive and important in African contexts, including power dynamics, the relationships between people, the policies and priorities for development, different ideas about what 'success' is and how it can be measured, and the balance between the rights of individuals, societies and nature".[3]

The rigours and constraints imposed on government-led M&E processes are inherently technocratic rather than people-centred. Consequently, M&E processes are often described as mandatory, compliance-focused, extractive, burdensome, and costly. This chapter argues that such terms would be less frequently used if conventional M&E methodologies became more inclusive, equity-driven, and participatory. This necessitates transforming the discourse surrounding the conceptualisation and utilisation of M&E by practitioners, moving beyond mere data collection and information processing, and embracing a comprehensive commitment to equity and inclusion. In the African evaluation context, evaluators must recognise that development requires a complex systems-informed evaluation approach that integrates sociocultural, economic, political, technological, and environmental factors, connecting local contexts with the global stage and Africa with the world (AfrEA).

Many reasons drive the instinct to actively improve and strengthen government monitoring and evaluation systems. Fiscal pressures and the increasing expectations of ordinary citizens continuously motivate governments to expand services and

2   Jara Dean-Coffey. 2021. Positionality posted on the Equitable Evaluation Initiative website at https://www.equitableeval.org/blog-main/positionality

3   African Evaluation Association. 2021. The African Evaluation Principles at https://afrea.org/AEP/new/The-African-Evaluation-Principles.pdf

enhance quality standards (Mackay 2007). This includes addressing the expectations from vulnerable, marginalised, and underserved citizens in an equitable manner. However, achieving this may be challenging as vulnerability and marginalisation within societies often stem from structural inequities, encompassing political, economic, and social dimensions. Without understanding and measuring these factors comprehensively, governments face difficulties in effectively addressing them. Additionally, when dealing with broad concepts such as poverty, inequality, and marginalisation, the diverse and nuanced aspects of people's lived experiences are often overlooked. Generalised, insufficient, outdated, or poorly disaggregated data, as well as data gaps, can lead to service gaps, leaving segments of the population unaccounted for in national statistics. Consequently, the lack of evidence negatively impacts the state's ability to meet the needs of the most vulnerable and marginalised groups in society, weakening its capacity for equitable service delivery.

The Changing Landscape of Government Monitoring Systems

Practitioners generally face the challenge of being unaware of critical equity and inclusion-related data. Holmes (2020:1) highlights the concept of "positionality", which refers to an individual's worldview and the position they adopt regarding a research task and its social and political context. The concept of positionality challenges the notion that the practice of M&E should be objective and free from values. Consider two government M&E officials working with police crime data and engaging with statistics related to gender-based violence (GBV), including rape, sexual assault, and femicide. Each official brings their professional and personal perspectives to the data. One may possess knowledge, empathy, and responsiveness to gender-based inequalities, while the other may hold a patriarchal understanding of gender roles. However, these perspectives can converge when acknowledging that each data point represents a complex individual with unique circumstances in their lived reality. This understanding allows these practitioners to shape the evidence to inform equity-responsive analysis effectively.

Equity-driven monitoring systems require practitioners who understand and are committed to using evidence to measure both the drivers and consequences of inequalities. Achieving this necessitates capturing evidence that reveals structural issues (drivers) and the socioeconomic impacts (consequences) of inequalities within societies, considering how they affect different segments of the population. In the case of police data on GBV, combining disaggregated data sets can provide a more detailed understanding of the victims and survivors, encompassing factors such as age, ethnicity, geographical location, marital status, and education level to a broader degree. At a more granular level, additional information such as social and economic

status, level of education, sexuality, and HIV status can be considered.[4] Working with a more comprehensive cross-section of data allows practitioners to identify patterns and intersections of inequity and exclusion. This, in turn, enhances their ability to provide evidence that informs the work of policymakers and programme designers. The Organisation for Economic Cooperation and Development (OECD 2018:72) emphasises the importance of national statistical systems pursuing sophisticated data disaggregation strategies. Current statistics typically capture national averages but fail to reveal disparities at the subnational, community, household, and individual levels. AfrEA advocates for evaluations that serve an equity purpose, recognising their potential as transformative exercises. They further emphasise that financing, commissioning, conducting, and utilising evaluations in Africa are "highly responsible tasks, especially when dealing with vulnerable communities and economies, developing institutions, and the rich diversity of worldviews, experiences, and traditions that define African societies" (AfrEA).

Monitoring systems that are designed to produce equity-informed evidence are most effectively managed by practitioners who comprehend the importance of moving beyond routine or standardised measurement processes and utilising the broader potentiality of data. The pervasive challenge of poverty in Africa and the way in which it is measured can be examined. In the development context, poverty is generally seen as the overarching manifestation of inequality. This is highlighted by the fact that poverty and inequality feature prominently in the United Nations' seventeen Sustainable Development Goals (SDGs).[5] In the broadest terms, the most widely used measure of inequality is the Gini coefficient, which ranges from zero (perfect equality) to one (perfect inequality). The Gini Coefficient is a generalised, population-based metric. It is not easily broken down to reveal the sources and consequences or the sections of the population most affected by inequality. Inequalities and disadvantages are embedded in the power dynamics of social structures such as class systems, cultures, religions, and gender relations, and often manifest in social institutions and socioeconomic systems. A more composite and multidimensional response is required to better understand the complexity and multidimensionality of the drivers of poverty and inequality, given that narrowing gaps in one area may not be sufficient to reduce disparities in other domains of well-being (Rohwerder 2016). Statistics South Africa (StatsSA) has undertaken work on a nuanced approach to the consideration of the multidimensional and intersectional forms of poverty to

---

4   The concept of "granularity" in M&E refers to the level of detail of data, based on the understanding that the more that detailed that data is, the more precise any analysis can be.

5   The SDGs were adopted by the United Nations in 2015 as a universal call to action to end poverty, protect the planet, and ensure that by 2030 all people enjoy peace and prosperity – see https://www.undp.org/sustainable-development-goals

develop a multidimensional poverty index, intended to move away from the blunt measurement of poverty, such as the World Bank's $1.90 a day threshold.[6]

An added layer of complexity involves the descriptors associated with the concept of inequality, such as vulnerability, marginalisation, disadvantage, deprivation, disempowerment, and social exclusion. Bok (2018) argues that social exclusion can be described as a comprehensive, multidimensional, and dynamic concept that generally refers to the limited opportunities of individuals to participate financially, socially, culturally, and politically in their societies. Khan et al. (2015) suggest that social exclusion is a process by which certain groups are systematically disadvantaged due to discrimination based on their ethnicity, race, religion, sexual orientation, caste, descent, gender, age, disability, HIV status, migrant status, or area of residence. It further refers to the processes behind the accumulated vulnerability and weakening of social rights, including discrimination embedded in public institutions, such as the legal system or education and health services, and social institutions like religion, cultures, workplaces, family networks, and the household. Sen, Kessler and Loveridge (2018) proposes a capabilities approach and argues that social arrangements should be evaluated according to the extent of freedom with which people can promote or achieve functionings that they value, suggesting that wellbeing should be measured according to what individuals can do (capabilities) as opposed to what they do (functionings). This poses a challenge for M&E practitioners who need to understand the assumptions made by Sen et al. (2018), and the way in which to develop tools that can measure capabilities.

Applying an Equity Lens to Monitoring Systems

An equity perspective makes the production and use of data fairer, more robust, and more accurate in general. Moreover, to ensure equity in any analysis process, the data being used must reflect the fact that an individual's experiences are not unidimensional (poor, illiterate, or HIV positive) but are based on multiple and intersecting dimensions, identities, and experiences. As an example, the effects of poverty, marginalisation, disability, and sexuality would intersect within the reality of a person who is a poor, rural, disabled lesbian of ethnic minority. Different combinations of demographics and identities create different types and experiences of inequality for various people. From an equity point of view, it is essential to gear monitoring systems toward generating data and analysis that reflects these very granular realities. Population-level data for disability (2.6% of the population) does not reflect different kinds or experiences of disability. Neither does it reflect the different levels of opportunity and access that disabled persons from different backgrounds may have. When monitoring systems work with blunt data, they tend to overlook or conceal the

---

6   The World Bank data portal at https://data.worldbank.org/indicator/SI.POV.DDAY

different socioeconomic realities of population sub-groups and, consequently, reduce analysts' ability to dissect issues of diversity, intersectionality, and difference. This lack of multidimensional data could impact policymaking where policies are designed based on inadequate or insufficiently differentiated evidence.

Figure 2.2:   *Wheel of intersectionality, Identiversity, [1 July, 2023]*

Common drivers of inequality relate to a lack of inclusive growth, lack of investments in human capital, lack of pro-poor fiscal policies and redistribution, lack of access to essential services and human rights, and a lack of political will to tackle the root causes of discrimination, as well as structural exclusions based on social, political, cultural, and economic factors. Every country has its unique systemic challenges in this regard, but these dynamics often play a role in perpetuating social, economic, and political inequities within societies. Issues related to legal discrimination, social expectations regarding gender and sexual identity roles, restrictions on bodily integrity, class and ethnic forms of discrimination, and various forms of exploitation and marginalisation, largely explain the persistence of unequal outcomes in employment, entrepreneurship, health and wellbeing, access to opportunities, and political representation.

With women making up just over 50% of the world population, gender is perhaps the most contentious area of inequality. The Organisation for Economic Cooperation and Development (OECD) notes that discrimination in social institutions, such as education, health provision, the labour force, and financial access, contributes to

obstacles in gender equality in development outcomes across all world regions.[7] However, the commitment to the implementation of equity-based monitoring and evaluation systems remains rhetorical, unless M&E systems within public services are able to develop and implement indicators that measure progress on achieving gender equality by understanding that the lived experience of women and girls is complex, heterogenous, and multidimensional.

The term 'institutional', as referred to by the OECD, speaks to the formal and informal systems, rules, and norms that structure and govern the social order and that obstruct or exclude people from social service provisioning, public employment, or other areas of social interaction (Fischer 2011). The parameters for acceptable decisions, choices, and behaviour for women and other marginalised groups in society are established by social institutions, and consequently define their roles and impact their life outcomes. These are undoubtedly challenging areas to measure if conventional M&E tools are used. Religious and cultural norms regarding sexual and reproductive health typically oppose notions of equity and rights, resulting in the disruption of equity-informed evidence for M&E practitioners to work with. The data challenges associated with an equity-oriented approach to managing data on abortion, contraception, age of consent, sex trafficking, cyber-grooming of underage girls, and female genital mutilation, should be considered. Gender-based discrimination in rights, opportunities, and outcomes interconnect and overlap, thereby further reinforcing women and girls' marginalisation.

Discriminatory laws, norms, and practices as measured by the OECD's Social Institutions and Gender Index (SIGI) and the African Development Bank's (AfDB) Africa Gender Index (AGI), map the limitations placed on the role of women as sexual beings, workers, entrepreneurs, healthy citizens, leaders at the national and local levels, and actors of human development. Based on these kinds of indices, higher levels of gender-based discrimination in social institutions are associated with lower equality in outcomes. Simplified, this means that countries with higher levels of discrimination in social institutions are further from achieving gender parity. In its 2020 report, the AfDB notes with emphasis that the limited availability of policy-relevant gender statistics poses a challenge to the inclusion of several fundamental aspects of gender inequality in the AGI.[8] The recognition that national and regional statistics are lacking in critical areas where exclusion and marginalisation are most pronounced is a positive step forward for advocates of a more equity-driven monitoring practice that integrates this approach into monitoring systems.

Fisher (2011) argues that exclusion is a pressing development concern. An equity-driven monitoring approach highlights the intersecting processes of exclusion which

---

7   OECD at https://www.oecd.org/dev/development-posts-gender-discrimination.htm

8   Africa Gender Index: Methodological and Statistical Report 2020:7.

are not effectively captured by poverty, inequality evidence, and methods of analysis, particularly those within contexts of high or rising inequality. Absolute and relative indicators often reveal little about processes of exclusion and marginalisation, and if they do, they tend to do so by revealing clues about the spaces within which exclusionary practices and processes might operate. It is understood that poverty is a problem statistically, but the various ways in which different "poor" or "marginalised" people are impacted is not as clear. Standard statistical sampling methods based on generic outcome indicators are the customary practice in the government and development sectors but are poorly suited for approaches that focus on equity or social inclusion. An example would be the way in which a blunt statistic, such as the HIV+ percentage of a population (usually disaggregated by age and gender), fails to capture differential experiences of access to testing, treatment, care, and support, as well as community and internalised stigma. To this end, more inductive methods, such as gathering granular data at the ground level, would better serve monitoring systems that can trace the kinds of implicit and explicit socioeconomic dynamics that affect people's lives. This would include interdisciplinary analyses of structural and institutional disjunctures and asymmetries operating across social hierarchies and among comparable cohorts within a social hierarchy, such as those with similar levels of educational achievements and employment expectations.

Reimagining the Monitoring Paradigm

The utilisation of the tools of equity, inclusion, and intersectionality requires a shift in methodological approaches to data collection and analysis and an epistemological shift in terms of what constitutes valid and recognised data. Part of the epistemological shift has been the movement from hard, empirical (quantitative) data to a growing interest in the more qualitative dimensions of intersectionality, which form an integral analytical layer for a better understanding of the complexity of inequalities and social exclusion. Originally posited by legal scholar Kimberlé Crenshaw in the field of critical legal studies, the concept of intersectionality aimed to deconstruct the application of laws utilised in legal cases to illustrate the ways in which the structures of law and society could be intrinsically racist. Crenshaw (1989:167) noted that intersectionality is a strategic way to place those who are currently marginalised at the centre of the conversation and is "the most effective way to resist efforts to compartmentalise experiences and undermine potential collective action". The concept of intersectionality as a framing construct has subsequently been applied in various disciplines, including radical sociology, feminism, gender studies, queer theory, equity and diversity studies, and critical race theory.

The concept of intersectionality has become a contentious term in the lexicon of conservative politics for its work in deconstructing racist, colonial, sexist, and homophobic discourse. Despite resistance, the concept of intersectionality has become a

valuable analytical tool throughout the last decade, particularly for those working on issues related to socioeconomic inequalities and marginalised or vulnerable populations within societies. While it has been broadly applied in qualitative research studies, it has only recently experienced a surge in quantitative research, owing to the technical and cost challenges associated with collecting more granular quantitative evidence. However, the implementation of a theoretical intersectionality framework into quantitative data analyses is gaining increasing interest in health research due to the understanding that the complex causes and mechanisms leading to health inequalities can be improved by the integration of an intersectionality framework (Mena et al. 2019). This can be similarly applied to other social sectors such as education, criminal justice, and social protection, where inequities constitute a significant determinant of access to fair and appropriate services. It is evident that people may continue to 'fall through the cracks' in policymaking and service delivery due to the lack of data to evidence and understand the unique challenges experienced by people facing multiple forms of inequality.

Many countries have constitutions that enable states to prioritise human rights, inclusion, and equity for all citizens, and rights-based pieces of legislation and policy frameworks that stem from crucial constitutional mandates. Some countries retain policies and legislation that undermine efforts to achieve equity and inclusion despite having a rights-based constitutional mandate. Challenges arise when government sectors are expected to realise different kinds of equity imperatives in their work or services, with these being dependent on existing policies as well as the quality of data available to policymakers and planners. In budget-constrained countries where public services are often inadequate, many citizens slip through the inclusion net due to their specific vulnerabilities not being captured in official data sets. Individuals who are at the intersections of disadvantages may struggle to have their needs met when "policies are developed using a single-factor lens, activated by single-factor trigger points, and/or developed to offer single-factor interventions" (Corus et al. 2016); they are "invisible" to official recognition. Crenshaw, and other feminists, point to the intersection of race and gender in countries such as the United States and Brazil, where women may face exclusion from jobs deemed more appropriate for men due to their sex. Women may be excluded from jobs considered "women's jobs" because of their race (AWID 2004). As a result, women of ethnic minorities specifically face exclusion from employment opportunities. If official data collection and analysis processes do not recognise these experiences, there is little likelihood that they will be addressed in any viable policy or programmatic way. It is evidenced by these methodological challenges that there is enormous scope for M&E and knowledge management practitioners, who work at the fulcrum between data generation and data analysis, to interrogate their practice and start applying new methodologies and tools to sharpen an equity-informed intersectional practice.

There exists a tendency within government to foreground the importance of stakeholder consultation and community engagements to foster understanding regarding people's lived realities, but these engagements can be formulaic and are incapable of allowing communities to feed local experience and knowledge into government monitoring systems. This highlights a lost opportunity, as government responsibilities (including the provision of healthcare, education, and social assistance) require ever more sophisticated evidence sources to enable the design of policies and programmes that are as inclusive as possible and are informed by the actual needs of people. The tendency of people being omitted from data is particularly distressing, as government policies and programmes that produce administrative data can inadvertently contribute directly to creating, enabling, and sustaining institutional and structural forms of discrimination and marginalisation. The desirable condition is one that allows cross-sector data sharing and integration, which enables the transformation of individual-level information into actionable intelligence for M&E practitioners. Ideally, more granular evidence can be used to understand urgent and long-term community needs and can result in the implementation of improved and more equitable services, systems, and practices that, according to the UN concept, "leave no one behind".

Reflections on the Discourse of M&E

Considering the reflections on what could be considered a "desirable state" for an equity-informed M&E practice, it is crucial to acknowledge the ideological challenge posed in terms of the implementation of an equity-informed monitoring process. Concepts such as human rights, social inclusion, and equity are socially constructed and hinge on a consensus regarding the definition of 'normal' in each society. As an example, 2021 saw 31 countries throughout the world legally recognise same-sex marriage, while 71 countries criminalise same-sex sexual activity, and 11 of those countries impose the death penalty in response to it.[9] Defining the concept of 'normal', 'acceptable', or 'decent' is not a simple task in many developing countries, where many people do not enjoy what may be considered a decent standard of living, due to the perspectives that countries or regions may have on critical developmental challenges. People who find themselves outside of their specific societal definition of 'normal' may be excluded based on their race, caste, or gender, due to social exclusion being structured around hierarchy. It is often difficult to 'objectively' identify those being socially excluded, as it is a matter of assumptions made, criteria adopted, and judgements used, which results in the concept of social inclusion and exclusion being contested. Equity and inclusion concepts focus attention on two central elements of deprivation: its multidimensionality and the processes and social

_____
9   Human Rights Campaign. 2022. Marriage Equality Around the World https://www.hrc.org/resources/marriage-equality-around-the-world

relations that underlie it. Social exclusion overlaps with poverty but extends beyond it by explicitly embracing the relational and distributional aspects of poverty. It is crucial to understand that M&E practitioners work in accordance with their position within a society and its values, in an evidence landscape that is fixed (i.e. global indicators) but fluid (i.e. the domestication of global indicators).

Work within a monitoring environment can be experienced as a one-dimensional technocratic exercise, with it being a mechanistic and functional "tick box" process of managing data inflows and assessing progress against indicators in performance frameworks. Every piece of human data tells a story or frames a real, lived experience. In viewing ourselves as development M&E practitioners, we can interact with data in ways that detail lived realities in a world inundated by inequality on multiple levels. Following Shanker (2018), the equity-driven M&E practitioner can reflect on and pose challenging questions about data, such as how a data point categorising a person as "unemployed" is informed by how that individual is racialised, classed, gendered, and sexualised. Having the necessary granularity of evidence to do these kinds of equity-informed analyses becomes critical in driving a more ethical and developmental practice.

With some issues that inform an equity-oriented M&E approach being outlined previously, the following consideration focuses on the means by which these approaches can be translated into reimagined practices. Efforts to introduce equity as a key monitoring component are relatively recent and have yet to be systematically mainstreamed into government monitoring and evaluation processes. Much of the innovation and experimentation in this regard is being tested in the development and philanthropic sectors, and to some extent in the private sector. While the theoretical underpinnings of intersectionality provide an essential lens for understanding social exclusion challenges, the M&E sector needs to grapple with the "how to do equity" conundrum. The focus within an African context lies in the move to "decolonise" evaluation, which entails a process of comprehensive review of the ideological and epistemological underpinnings of current practices and knowledge. These should be gathered, taken apart, broken up, and critically examined to identify their benefits and weaknesses in terms of equity-informed practice. This is an emerging area of interest within the M&E field, and it represents opportunities for technical experimentation and innovation. Bowleg (2008), for example, argues that a key dilemma for researchers and M&E practitioners working with intersectionality is to grapple with the assumptions inherent in measurement as well as qualitative and quantitative data analyses, and to transform them so that equity can be integrated into practice.

In the M&E environment, an inherent bias exists towards easily quantifiable, pragmatic, and reductionist data that simplifies the process of making sense of a complex world. However, this approach contradicts the central tenet of intersectionality, which is that social identities and inequality are interdependent for many groups of people and are not mutually exclusive. Programme interventions and their associated M&E frameworks rely on output and outcome level indicators to measure progress and do not attempt the building of intersectionality thinking into the process. Discrete and continuous variables, such as gender, race, ethnicity, and age, are measured as independent factors but attempts to understand how these variables intersect and interact with each other are rarely undertaken. Considering this gap, interpretation becomes one of the most substantial tools in the methodological toolbox of researchers, data analysts, and those working with M&E systems. Significant progress is being made, for instance, in the United Kingdom, where charters have been put in place that require government institutions across the board to collect equity and inclusion data that ethically reflects intersectionality considerations.

Equity-oriented practitioners across the policy, programming, monitoring, analysis, and reporting spectrums can be sensitised and informed by a range of approaches such as heuristics, assessment tools, checklists, and matrices. Second and third-wave feminists and queer theorists working in gender and sexuality studies carried out much of the initial analytical thinking regarding embedded socioeconomic and political inequity. One such theorist, Caroline Moser, introduced the Gender Analysis Framework in the 1980s as a means of supporting an integrated gender-planning perspective in all development work, concentrating on the power relations between men and women. This approach allowed planners to engage with the complexity of inequality and introduced the subordination of women into planning and monitoring discourse.[10] Some of this analysis has been translated into public and development theory and practice. The gender-responsive assessment scale, a tool developed by gender studies theorists, enabled policymakers, programme designers, and M&E practitioners to reflect on and assess the level of gender-responsiveness manifested policies, strategies, and programmes. This also resulted in the interrogation of the kinds of indicators being used to collect critical data. The value of these analytical tools lies in their ability to be adapted to the assessment of gender-responsiveness as well as other areas of inequality, complimented by an additional intersectionality lens. The following scale is adapted from the World Health Organisation (WHO) Gender Responsive Assessment Scale, which was developed for managers in the health sector:[11]

---

10   European Institute for Gender Equality. 2019 available at https://eige.europa.eu/sites/default/files/mh0319271enn_002.pdf

11   https://www.who.int/gender/mainstreaming/GMH_Participant_GenderAssessmentScale.pdf

Table 2.1:   *WHO, Gender Responsive Assessment Scale*

Equity Responsive Assessment level	Consequences for policies, strategies, and programmes	The extent to which an Intersectionality Lens is applied
Equity-Unequal	• Perpetuates inequality by reinforcing unbalanced norms, roles, and relations. • Privileges one or more groups over others. • Often leads to one group enjoying more rights or opportunities than the other.	• There is no understanding of differences within groups or how social norms determine these differences. • Groups are homogenised and essentialised.
Equity-Blind	• Ignores societal norms, roles, and relations (gender, religion, culture, etc.). • Very often reinforces identity-based discrimination. • Ignores differences in opportunities and resource allocation for different groups/populations. • Often constructed based on the principle of being "fair" by treating everyone the same.	• Poorly informed understanding of different identities within groups is not accounted for. • Intersections of identity and experiences are not seen. • Different identities and experiences are collapsed into generalised categories.
Equity-Sensitive	• Considers societal norms, roles, and relations. • Does not address inequality generated by unequal norms, roles, or relations. • Indicates equity awareness, although often no remedial action is developed.	• Some understanding of how social norms construct identities. • Some understanding of how intersecting identities determine people's life experiences. • Limited understanding of how to respond to intersectionality at a policy or programme level.

Equity Responsive Assessment level	Consequences for policies, strategies, and programmes	The extent to which an Intersectionality Lens is applied
Equity-Specific	• Considers societal norms, roles, and relations for different groups and how they affect access to and control over resources. • Considers the specific needs of different groups. • Intentionally targets and benefits a specific group of people to achieve certain policy or programme goals or meet certain needs. • Makes it easier for people to fulfil duties that are ascribed to them based on their perceived socioeconomic roles.	• A more nuanced understanding of how intersecting identities determine access to and control over resources. • Consideration is given to complex identity formation within groups. • Some ability to respond to intersectionality in the formulation of policies and programmes.
Equity-Transformative	• Considers gender norms, roles, and relations for different groups and that these affect access to and control over resources. • Considers the specific needs of different groups. • Addresses the causes of inequities. • Includes ways to transform harmful social norms, roles, and relations. • The objective is often to promote equality. • Includes strategies to foster progressive changes in power relationships between different groups of people.	• A strong theoretical grasp of intersecting identities and differential outcomes within and between groups. • Policy/programme responses consider intersectional impacts. • Addresses the causes of inequities at a granular level. • Includes ways to transform the modalities of intersectionality that negatively impact people's life experiences and outcomes.

An assessment scale such as this needs to be utilised across the policy and programme cycle to be effective. A challenge often faced by equity-responsive monitoring practitioners is the immensely complicated means of measuring the performance of strategies and programmes which have failed to integrate equity and intersectionality considerations during the conceptualisation and design phase. However, monitoring systems can be proactive in inserting equity adaptations from the bottom up. A recent application is in the public health field of HIV/AIDS and

human rights, where the Global Fund (2019) has, in its "Breaking Down Barriers" initiative, set out programming parameters to address inequities in access to critical health and legal services.[12] In many countries in Southern Africa, HIV has become a manageable health condition. However, prevalence remains high in vulnerable populations, including the LGBTQIA+ community, sex workers, drug users, migrant communities, as well as vulnerable girls and women, where stigma and discrimination can impede access to public health and legal services. This example illustrates how the consideration of modifiable societal and contextual factors in the reduction of health inequalities becomes increasingly important both to implement and measure.

From the data generation and monitoring perspective, the challenge is that many aspects of individual lives remain largely unaccounted for in public data sources such as national censuses, household surveys, and demographic and health surveys. In the case of both communicable diseases (HIV, TB, and COVID-19) and non-communicable diseases (obesity, alcohol use disorder, and mental health) there are many reasons for this disparity which lie outside of the control of M&E practitioners. Some of these include the criminalisation of same-sex relations, sex work, drug use, migrancy, and their associated levels of societal stigma, discrimination, and marginalisation. Without reliable disaggregated data that reflects the intersectional experiences of people's lives, the challenge for policymakers and programmers lies in the development of targeted interventions that are appropriately evidence-informed and are designed to reduce or eliminate inequities in access and opportunity.

Moving to More Equitable, Intersectional, and Inclusive Measurement

A challenge faced by monitoring systems includes the collection of data from populations that are difficult to reach. These populations often view the government as a threat to their safety and well-being, making this problem particularly relevant in terms of the government data collection processes. Civil society organisations that specialise in working with vulnerable populations are taking innovative strides in data collection from these populations. One such organisation, Frontline AIDS, has developed a peer-driven data collection system called REAct which monitors human rights violations and focuses on HIV programming as well as advocacy for marginalised people who are vulnerable to or affected by HIV and AIDS.[13] Academic research institutions may also undertake Integrated Biological and Behavioural

---

12   The Global Fund. Human Rights https://www.theglobalfund.org/en/human-rights/#:~:text=Through%20our%20Breaking%20Down%20Barriers,HIV%2C%20TB%20and%20malaria%20services.

13   The technicalities of the REAct system are set out in the REAct User Guide https://frontlineaids.org/resources/react-user-guide/

Surveys (IBBS) that use peer researchers to collect personal data from hard-to-reach populations such as at-risk ethnic minorities, transgender people, sex workers, and survivors of rape. These approaches effectively use bespoke data collection tools to collect data that accurately reflects the intersectional reality of individuals who report into the system and makes referrals and tailor-made responses easier. The REAct system, for example, is run by peer community health workers and paralegals accepted and trusted by vulnerable people. They can collect data that official data collection methods cannot and feed it into national data collection systems such as health information systems. These applications in health research mark an emerging research paradigm that seeks to move beyond single or typically favoured categories of analysis (i.e., sex, gender, race, and class) to consider simultaneous interactions between different aspects of social identity, as well as the impact of systems and processes of oppression and domination (Palmen 2021).

The intersectionality wheel shown in Figure 3 (WHO 2020) is a useful heuristic that enables practitioners to think about what intersectionality means in practice. It uses a variant of the socioecological model that is premised on the overlapping and intersecting domains of the individual, the community, the society, and the structural dimensions of that society.

Figure 2.3:   *World Health Organisation, Intersectionality Wheel, s.a. (World Health Organisation 2020)*

The intersectionality wheel illustrates how multiple individual characteristics (such as age, gender, and education) interact within broader processes of social discrimination (such as ableism, sexism, and racism) and structural barriers (such as political, class, and economic) to shape an individual's position within a given society. It allows researchers to model demographic characteristics across different levels by capturing people's lived experiences within the intersecting macro, meso, and microsystems. Intersectional approaches seek to consider the positions of all members of a given society and aim to illuminate the position of the most marginalised within existing power dynamics through the lens of the barriers faced in accessing equitable public services.

Government agencies in certain parts of the world actively use intersectionality as a monitoring strategy by addressing identity stereotypes and individuals' unique and intersecting experiences that result from race, gender, sexuality, disability, or other forms of vulnerability. Women and Gender Equality Canada has developed a process called Gender-based Analysis Plus (GBA+), which is an analytical process that can be used to analyse the "gendered" aspects of Canadian government policy to assess the different experiences of women, men, and non-binary people within policies, programmes, and initiatives.[14] The 'Plus' component considers the intersectionality of identity factors and how the relationships between these identity factors impact the way in which government programmes and initiatives are designed, monitored, and experienced. The Ontario Human Rights Commission in Canada argues that intersectional analysis should become one of the lenses through which the social context of the individual can be examined and, in some measure, should assist in addressing social conditions relating to poverty, low income, and homelessness through improved data (OHRC 2021). In countries such as Canada, the United States, New Zealand, Australia, Mexico, and South Africa, First Nation people are often the most marginalised and stigmatised population segments.

Developed in the United Kingdom in 2005, the Athena Swan Charter is a globally used framework used to support and transform gender equality within higher education (HE) and research. Initially established to encourage and recognise a commitment to the advancement of the careers of women in science, technology, engineering, mathematics, and medicine (STEMM) employment, it is now being used in many countries to address gender and identity equality more broadly, no longer focusing solely on progression that affects women (AdvanceHE 2021). The South African Workplace Equality Index (SAWEI) is a similar monitoring system that enables South African companies to measure their progress against independently determined and research-based best practices regarding LGBTQIA+ inclusion in the workplace. This system uses a survey consisting of six elements that measure equity in the structures

---

14 Government of Canada. Women and Gender Equality Canada available at https://women-gender-equality.canada.ca/en.html

and behaviours of companies.[15] The Gender Lens Investing Initiative is an intervention that stems from the Global Impact Investing Network (GIIN) which supports private sector impact investors in the active integration of a gender lens strategy into their investment portfolios to address the systemic issue of a lack of gender equity that has been present in the business and investment community throughout its existence.[16] Part of this initiative has been to set up monitoring systems that can track the impact of investments on the lives of women and girls. The International Foundation for Electoral Systems (IFES) has developed a new assessment framework to identify intersectional barriers and opportunities related to people's political participation with multiple social identities, including gender, disability, and age.[17] This participatory tool was developed based on the recognition that marginalised populations, including women, persons with disabilities, LGBTQIA+ persons, and people from ethnic or religious minorities, experience systemic discrimination in many countries. In some cases, victims of physical or psychological violence have had their political and constitutional rights undermined by voting rights barriers (Atkinson 2018).

Although marginalised groups experience well-documented barriers to political participation, there remains little analysis of the intersecting challenges or common goals between individuals in each area. The M&E field has recently undergone growth through the emergence of The Equitable Evaluation Initiative (EEI), which has assisted in the advancement of equity and the expansion of rigorous notions of validity and has fuelled the embracing of complexity.[18] This approach shifts the current evaluation paradigm and encourages mindsets and practices to evolve towards a more equity-oriented practice. Its Equitable Evaluation Framework supports M&E practitioners working with Foundations and Civil Society Organisations to replace conventional narratives that marginalise, minimise, and disrespect people with ones that reflect an understanding of the systemic and structural barriers that limit the possibilities of people and their ability to thrive. These methodologies provide equity-focused tools for data collection that allow people who identify with multiple social identities to share the different ways in which they participate civically and politically, prioritise identified solutions to barriers, and share their opinions in targeted surveys. Where used purposely, these methodologies are a critical first step

---

15   The SAWEI is managed by the LGBT+ Management Forum that works with companies across South Africa to create safe and equitable workplaces that enable lesbian, gay, bisexual, trans (LGBT+) professionals to contribute to their fullest potential, and the index reports are available at http://lgbtforum.org/news/view/sawei-2021-results-announced-with-5-gold-tiered-companies

16   Global Impact Investing Network,. Gender Lens Investing Initiative https://thegiin.org/gender-lens-investing-initiative

17   International Foundation for Electoral Systems. Participation and Inclusion https://www.ifes.org/issues/participation-and-inclusion

18   Equitable Evaluation Initiative https://www.equitableeval.org/

towards generating intersectional data that can feed into and augment other more standardised datasets. UNESCO's World Inequality Database on Education (WIDE) is another useful interactive tool for illustrating how different forms of disparity such as gender, household wealth, ethnicity, religion, and residence, play an important part in shaping opportunities for and access to education.[19] It visually demonstrates how overlapping disadvantages can compound education disparities. The World Inequality Database was initially created as the The World Top Incomes Database (WTID) in January 2011 with the aim of providing free and convenient access to all the existing series. It then expanded to include a series on income inequality for more than 30 countries which spanned over most of the 20th and early 21st centuries, with over 40 additional countries now under study.

Applying an intersectional approach (or lens) helps assess the potential impacts of initiatives – positive or negative - based on their multiple identity factors, enabling programmers to identify risks and potential challenges early and create mitigation strategies in response. An intersectional approach should be applied at all stages of an initiative, from development through implementation and monitoring to evaluation. The consideration and identification of people's diversity and multiple identity factors assist practitioners with the innovation and consideration of issues and policies in different, more reflective ways. This has clear implications for the design of programme monitoring systems. Most M&E systems, whether in government, the development sector, or civil society, are generally set up for administrative compliance, accountability, and routine reporting purposes, and tend to focus on broad policy and programme outputs that are easily measurable and reportable. This is crucial for government accountability to the type, scale, cost, and reach of their services to citizens. However, in the developing world context, most public services must address significant socioeconomic disparities at a broad societal level, such as addressing poverty, unemployment, and inequality. In most instances, their monitoring systems are not optimally equipped for tracking progress on gender and social inclusion and cannot, as a result, provide policymakers with data that is sufficiently granular and disaggregated. This could be a capacity issue, a design issue, a practice issue, or a combination of all three. In some cases, there may also be a lack of political appetite for collecting data on certain population segments.

For an effective monitoring system to be responsive to gender and social inclusion, the capacity, design, and practice issues need to be reappraised and reengineered. A more responsive monitoring system should provide more refined data to highlight the equity gaps and enable managers to better assess progress on key equity, gender, and social inclusion concerns, leading to more evidence-informed policy and programme decisions. There is also a need to have a more focused emphasis on advocacy and mainstreaming work with practitioners if there is a vision that

19   UNESCO https://learningportal.iiep.unesco.org/en/blog/one-way-of-measuring-marginalization-in-education

they can become developmental monitoring and evaluation specialists rather than M&E technicians. In reality, sourcing and producing standard data is not enough to make that data useful in terms of addressing equity considerations. Citizens are not interested in data; they are interested in improved services being designed and delivered based on available data and information that surfaces their realities. If that data is inadequate and only partially inclusive of a full range of citizen experiences, then the evidence bases for building, strengthening, and sustaining equitable services are compromised. Conventional monitoring systems within the government draw on available data ecosystems and produce compliance data or data that informs the indicators used for performance planning. The challenge lies in the data as well as M&E practitioners themselves being "blind" to the equity gaps in the data and the inadequacy of the data in informing an understanding of intersectionality issues. Efforts to address these equity gaps across the government policy and planning cycles are manifested in initiatives such as gender-responsive budgeting. This is a growing area of interest for developing country governments that are aware of the need to address gender inequalities across all sectors. In South Africa, for example, the Gender Responsive Planning, Budgeting, Monitoring, Evaluation, and Auditing Framework (GRPBMEA) is being championed by the National Treasury.

In some cases, grounds such as sex, race, ethnicity, or disability may intersect and produce unique effects creating "discrete and insular minorities" who are socially handicapped due to these same characteristics (OHCR 2016). Simple examples here include the situation of undocumented young migrant women, sex workers, or transgender women whom the state may deny critical sexual and reproductive health services due to their liminal status. Such situations may be typically thought of as people "falling through the cracks" or "being left behind" by government services that are intended to be inclusive and equitable. The lack of sufficiently disaggregated data (or no data at all) on these populations highlights the importance of monitoring systems of creating mechanisms that reflect and capture individual and group experiences based on multiple identities linked to more than one ground of discrimination. Alternatively, any one of these characteristics may intersect with other grounds such as social assistance, family status, and a further link to economic and social and class status to create unique experiences for the individuals that current developmental frameworks ignore. Even when combined with other grounds such as social assistance and family status, the extent of the discrimination may not be revealed by a traditional, non-intersectional approach.[20]

An intersectional analysis can be informed by developments in gender equality analysis, critical race analysis, disability rights analysis, and equality rights juris-prudence. These strategies have been developed to address the stereotypes and

---

20 Ontario Human Rights Commission https://www.ohrc.on.ca/en/intersectional-approach-discrimination-addressing-multiple-grounds-human-rights-claims/applying-intersectional-approach

unique and intersecting experiences of individuals because of race, gender, or disability. They would form a necessary part of the contextual and analytical framework. However, Runyan (2018) warns that a tokenistic approach could obscure the transformative intention of this approach and dilute the nature of inequalities without the impetus to use monitoring data to address the underlying structures that produce and sustain injustice. The relevance of this for monitoring systems is that a lack of critical equity data should be a concern for practitioners and management staff, as programmes with weak monitoring systems are not likely to be easily evaluable in terms of impact. Referring specifically to international development work Riddell (2014) argues that there is a lack of reliable, consistent, and robust programme information to answer whether development interventions work, partly because of inadequate baseline data and the often-weak monitoring upon which the assessment of impact is based.

A case in point is the stubborn persistence of gender-based violence (GBV), which is an increasingly visible societal challenge but is also one that yields concerningly little data. Accurate and reliable generalised data on the prevalence and incidence of GBV has been notoriously difficult to collect. Therefore, it is challenging to analyse available data for evidence-driven and equity-oriented programming. The intersectionality lens disrupts this problem further by calling for a greater understanding of how GBV impacts women and girls based not just on their gender but also on intersecting identities such as race, class, locality, and sexual identity. Based on this, much work is needed to make monitoring systems more inclusive and more responsive to equity considerations. A key consideration arising from the GBV challenge is that data "talks about" victims and survivors rather than "talks with" them. The means and extent to which practitioners decide to develop monitoring systems, including indicators, depend on their specific aims. They can simply collect data, such as police statistics, about a particularly concerning societal issue such as GBV, or they can go a step further by making sure that primary stakeholders, particularly those who are typically marginalised or excluded, participate in some way in to inform monitoring processes and practices so that their experiences and understandings feed directly into social transformation agendas.

Building an Equity-Oriented M&E Practice

This chapter has attempted to outline some of the theoretical underpinnings of an equity-oriented M&E practice and has pointed to some of the initiatives being undertaken worldwide to pilot innovative approaches. For those training, working with, or mentoring emerging M&E practitioners, several critical interventions can lead to an impact in strengthened interest in and commitment to the theory and praxis of equity-led monitoring systems.

Building a Commitment to Equity

As technical systems, monitoring systems will be as effective or ineffective as allowed by their design and the ability of their practitioners. Ideally, the design of monitoring systems should be informed by equity issues. Monitoring systems also need to be appropriately geared to tracking the routine outcomes of policy and strategy provisions on equity issues of national concern at both the macro and the micro level. Examples of these include gender equality issues related to remuneration, financial independence, political participation, safety and security, sexual and reproductive health, and bodily integrity. Programme designers and M&E specialists are generally well-attuned to equity issues in the development sector, as their work often focuses on vulnerable, marginalised, and at-risk populations. Even though government service delivery in developing countries carries an immense equity burden, public sector agencies often lack adequate sensitisation and awareness around effective integration of equity and intersectionality considerations into their work. Monitoring systems may be characterised by data gaps that result in the needs of specific population segments being unintentionally overlooked or intentionally ignored. As an example, many countries see ethnic and religious minorities, political activists, members of the LGBTQIA+ community, and sex workers being heavily stigmatised, marginalised, or even criminalised, and deliberately excluded from national statistics. Emerging practitioners must learn to understand how power dynamics include and exclude people and how this impacts the quality and availability of equity-informed data.

Awareness-Raising, Values Clarification, and Responsibilities as a Public Servant

Public services operate as significant sources of employment within countries and have extremely diverse workforces. The ideal scenario from a public service and administration perspective is to ensure that those working as public servants are inducted, capacitated, and assessed according to conditions of employment and prevailing policies and regulations. In the South African context, this requires that public servants at all levels comply with constitutional mandates on core human rights and equity issues and that equity is integrated into planning and programming initiatives.

Government in-service capacity development programmes on equity-related issues are well established, although their efficacy has not been thoroughly evaluated. In South Africa, the National School of Government runs a range of in-service training programmes for public servants, including ones that cover issues such as human rights and gender mainstreaming, but the impact of many of these programmes is yet to be evaluated. However, it has created a context for organically building on work that has previously been conducted within public services to mainstream

gender as standard practice across all sectors and disciplines of government. For M&E practitioners, this should arguably have resulted in a better understanding of how to work with sex-disaggregated data and undertake gender-responsive data analysis work. If this assumption is correct, these understandings can be extended to include an equity and inclusion practice that uses intersectionality as a core monitoring tool. This involves the sensitisation of M&E practitioners regarding the value of increasing equity dimensions in their data products by using intersectional analysis in a strategic and targeted way.

The result of this process brings about an additional challenge that sees individuals within the workforce lacking the knowledge, skills, and integrity to proactively address certain equity issues that form a part of their work mandate. An example of this involves cases of healthcare workers who possess personal value systems, life experiences, and cultural backgrounds which contradict their obligation to provide non-discriminatory and non-stigmatising services. In many countries, Ministries of Health provide values clarification training for healthcare workers to sensitise them to human rights issues, inclusive care, and equitable health service provision. These kinds of interventions give healthcare providers a chance to examine their values, attitudes, and beliefs and change attitudes that hinder the service progress, while honing professionalism among the health workers. Perhaps the most challenging example here is the constitutional right that healthcare providers have in countries such as the United States and South Africa to refuse to provide legally-sanctioned abortions based on freedom of conscience, religion, and opinion. The same concerns could apply to M&E practitioners who may already have a negative predisposition to collecting and analysing monitoring data relating to specific population groups or actions. Working with equity and intersectionality issues presupposes that the following preconditions should ideally be in place:

- Pre-service and in-service education and training to provide adequate knowledge and understanding of equity issues integral to effective and efficient monitoring systems.

- Strengthening a theoretical understanding of how data can be used to drive equity-informed monitoring systems.

- Values clarification training is provided as a standard practice to support staff in reconciling personal value systems with the requirements of their job responsibilities.

- M&E managers are sufficiently capacitated to provide leadership, guidance, and mentorship to staff in applying equity-focused monitoring systems and using intersectionality and inclusion frameworks.

Building Technical Expertise

Typically, monitoring systems work with linear and unidimensional data points, which do not necessarily provide insight into the complex ways that inequalities operate at the individual, societal, and structural levels. Therefore, focusing on identities is only a part of the work that an equity-oriented practice requires. A composite indicator, such as the one used to populate the UN's Human Development Index, ranks countries annually based on their level of human development and is informative but incomplete as it reflects the effects of inequalities rather than the causes thereof.[21] More difficult questions need to be addressed around structural inequality and the dynamics that create and maintain systems of identity-based inequality. This means understanding the drivers of inequality and the consequences of inequity in the real world for specific populations of people and how data does or does not evidence this. In the context of monitoring systems, practitioners need specific guidance and examples about how to import inclusion and intersectionality as analytical frameworks into their practice. This shift requires an express connection of individual experiences of discrimination with the structures of power and exclusion from which discrimination is bred. Reflexive consideration is required by practitioners of the systems and processes that they use, and how these may, by virtue of their design, perpetuate the use and application of equity-blind data.

An equity-informed framing of the monitoring practice can lead to a more analytical M&E practice that recognises groups of people as having multiple diverse and intersecting identities that impact how they understand and experience government initiatives and services. Public sector monitoring functions are normative and rule-based, providing routine evidence sets that feed into programme management processes. This iterative data collection process informs government decision-making in the context of policy-making programming and implementation. Data generated through statistical and M&E processes will routinely be disaggregated along a limited spectrum of variables, among other things gender, age, race, location, and employment status. A monitoring and evaluation officer coming into this environment will understand their role within the framework of such a normative, rule-based system and may not necessarily be aware of the wider potential of monitoring systems for driving an equity-informed practice.

Building technical capabilities to operationalise equity-informed practices should be integrated into pre-service and in-service education and training. This needs to happen in a very practical way, using approaches that have been designed in a range of different sectors to address equity considerations. Generally, researchers

---

21 The Human Development Index (HDI) is a summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable and have a decent standard of living. The HDI is the geometric mean of normalized indices for each of the three dimensions. bit.ly/3rVbwTT

have relied primarily on qualitative methods to investigate intersectionality, raising questions about the epistemological and methodological possibilities of generating and using quantitative intersectionality data. In this regard, practitioners could engage with studies such as the one published in the New England Journal of Medicine that explores how gender and race influenced the referral of patients for cardiac catheterisation (Shulman et al. 1999). They had data on both race and gender and produced a statistical analysis. They did a "main effects" analysis in which they looked at the influence of gender, followed by looking at the influence of race. They then combined these main effects additively and created a chart that looked like the one in Figure 4 seen below:

Figure 2.4:   *Bowleg and Bauer, Using a quantitative intersectionality-based model, 2016*

Bowleg and Bauer (2016) have used this example to point out that adding together main effects such as gender and race does not produce an intersectional analysis. A different set of results is found when the analysis is redone using a quantitative intersectionality-based model. It quickly becomes apparent that the actual bias here is against black women. The initial incorrect results showing the lower odds ratios for white women and black men in the additive model included bias being shown to black women. As Rouhani (2014) notes, this happens because quantitative researchers traditionally seek to address issues of social inequity by investigating axes of inequity, such as race, gender, class, and sexuality, and only considering the potential interconnectedness of these axes. To build an intersectional model,

practitioners need to move beyond an additive model or algorithm that is built, for example, as **Outcome = Race + Gender + Sexual Orientation**. This model can be used to understand the effects of one of the predictors, such as sexual orientation, on the outcome while holding other predictors constant. Holding gender constant while looking at the impact of race and sexual orientation does not, however, tell us whether the impact of these individual characteristics differ when they are allowed to fluctuate. Alternatively stated, what is the effect of sexual orientation when gender is allowed to be either male or female? This is the nature of core questions of intersectional analysis. Intersectionality posits that experiences at an intersection are co-constituted and must be considered jointly. This distinction between additive and intersectional approaches maps onto quantitative distinctions between main effects and heterogeneity of effect (Hancock 2007). By adding an interaction term (essentially, multiplication), the new model now looks like **Outcome = Race * Gender * Sexual Orientation**. This model can answer how the outcome changes for different combinations of the variables by looking at all three predictors together rather than individually. The multiplication in the model accurately estimates the simultaneous and layered effects of the different variables. In an intersectionality-informed analysis, the additive approach is incorporated as an initial 'baseline' model, upon which further analyses are applied using multiplicativity to account for the conditional effects of intersecting categories on a social outcome (Rouhani 2014).

A key principle of equity in data and intersectional analysis involves acknowledgment that the inclusion of individual-level data alone frequently produces biased and incorrect results. Without any context, data analysis can yield incorrect and skewed data outcomes. For example, we may have national or sub-national level demographic data that tells us about the number of unemployed young women in the 18-to-24-year cohort, disaggregated by race. However, it tells us little about these women beyond those core variables. The building of an intersectional model involves taking additional steps to look at variables and data that measure the context and communities in which individuals are situated. For instance, in a model about the effects of age, gender, and refugee status on educational outcomes, it is important to include measures of how accepting each community is to refugees. It could also include variables measuring the availability of education in various languages, systemic regulations on gender and education, local levels of xenophobia, and other potential forms of human rights violations. Various statistical methods enable this type of analysis, the most common of which are multilevel models designed to include variables measured at the individual level and several broader levels of aggregation, such as community and country (Sage 2020). Like traditional regression, these models can include multiplication, not just addition, as posited by the intersectionality paradigm.

In her Primer on intersectional analysis, Rouhani (2014) includes an example exploring the intersectional effects of race, education, and urban area. The different urban areas have different policies, so this variable acts as a measure of structural level influence in the model. Researchers could conduct cross-contextual comparisons that would evaluate the impact of this policy across urban settings to empirically investigate how policy constructs the relative power and privileges within the system. This could be done through a comparison on cumulative years before and after the policy introduction throughout cities and states that enacted the legislation versus those that did not.

Table 2.2:   *Rouhani, Applying an intersectional multiplicative approach, 2014*

		Odds Ratio			
		Chicago	Los Angeles	New York	San Francisco
Race X education interactions					
Black	OR less than high school	3.280	1.752	2.580	4.850
	OR high school graduate	1.896	1.200	1.950	2.540
	OR some college	1.461	1.390	1.428	2.102
	OR college graduate (ref)	1.000	1.000	1.000	1.000
White	OR less than high school	2.133	1.450	1.890	3.504
	OR high school graduate	1.621	1.325	1.450	2.320
	OR some college	1.181	1.120	1.320	1.950
	OR college graduate (ref)	1.000	1.000	1.000	1.000

There are packages and examples of multiplicative and multilevel models that can be used by M&E practitioners to add an intersectional analysis to their work which are available in most standard statistical packages. There is one example using R on multiplication or interactions as well as one on multilevel modelling. For SPSS users, there is one on interactions and one on multilevel modelling as well.

In reality, individuals are not single data points (e.g., male or female, literate or illiterate), and no individual lives one of their characteristics at a time. Our lived experiences are a simultaneous combination of all our characteristics and experiences. If we want to use data to reflect people's lived experiences more accurately, then we need to use it in combination with other data sets. This is a multiplicative approach and allows us to see how different characteristics are interacting or intersecting and then enables us to better analyse the positionality of an individual within a society. The use of multiplicative models allows practitioners

to see how different characteristics are interacting or intersecting (Krause 2021). When this frame is used as an equity-oriented monitoring tool, the data shows that cumulative marginalisation is much stronger than the marginalisation of one group at a time. When individuals are members of more than one marginalised group, the effect can be cumulative, as illustrated in Figure 5 below:

Figure 2.5:   *Krause, Applying the multiplicative approach to demonstrate cumulative marginalisation, 2021*

This kind of mapping demonstrates the effects of cumulative marginalisation. The more marginalised identities a person holds, the higher the order of intersection and the lower their probability of being treated equitably. Figure 5 looks at the cumulative marginality of women and the likelihood of their engagement in political activism. From the chart, it is possible to see that a woman has a probability of 17% regarding active involvement. The probability of a woman from the lowest social class getting involved is 9%. However, a woman from the lowest social class and belonging to the marginalised ethnic group has a 7% probability of getting involved. In this example, data is used multiplicatively instead of additively to highlight the lived experience of cumulative marginalisation. Using complex, multi-layered demographic data

in research and monitoring systems is not always simple. However, it is crucial to ensure that the work being done is equitable and that it gives practitioners a more granular understanding of the issues they are working with.

Conclusion

Arguably, M&E systems within government are slow to respond to trends and innovations around M&E within development, civil society, and academic sectors. The major players in the development field, including the World Bank, the regional banks, the OECD, the UN agencies, and international NGOs, are undoubtedly leaders in the development of equity-responsive monitoring and evaluation systems. In many cases, they work alongside government entities to strengthen monitoring systems in raising efficacy in the collection, processing, and utilisation of equity-informed data. Development agencies and other academic and civil society organisations working at the forefront of monitoring and evaluation innovation can only do so much. Governments have the mandate and the obligation to provide for their most vulnerable and marginalised populations through evidence informed and equity-driven policymaking and programming. Government M&E systems in Africa can also draw on cutting edge work that is being done through the auspices of organisations such as AfrEA and Twende Mbele on the development of indigenous and Afrocentric M&E practices that focus on equity as a core epistemological framework.

This chapter has attempted to provide reflections on how practitioners working in the public service can broaden their thinking about their practice and has suggested how routine monitoring systems can expand their remit by integrating methodologies for measuring equity, inclusion, and intersectionality.

Reference List

AdvanceHE. 2021. *The Athena Swan Charter* https://www.advance-he.ac.uk/equality-charters/athena-swan-charter

Agrawal, R., Kalugampitiya, A., Rinxin, J. and Hashim, K. 2017. Leveraging Monitoring and Evaluation Systems for Good Governance. *Evaluation for Agenda 2030: Providing Evidence on Progress and Sustainability*. Exeter, UK: IDEAS http://web.undp.org/evaluation/documents/Books/Evaluation_for_Agenda_2030.pdf

Atkinson, V. 2018. *New Intersectionality Assessment Framework. International Foundation for Electoral Systems* https://www.ifes.org/news/new-intersectionality-assessment-framework

Bok, K. 2018. Definitions and Measurement of Social Exclusion - a Conceptual and Methodological Review. *Advances in Applied Sociology*, Vol.8 No.5, 2018 https://www.scirp.org/journal/papercitationdetails.aspx?paperid=84529&JournalID=1002

Bowleg, L. 2008. *When Black + Lesbian + Woman ≠ Black Lesbian Woman: The Methodological Challenges of Qualitative and Quantitative Intersectionality Research. Sex Roles* 59, pp. 312–325 https://doi.org/10.1007/s11199-008-9400-z

Bowleg, L. and Bauer, G. 2016. Invited Reflection: Quantifying Intersectionality. *Psychology of Women Quarterly* 2016, Vol 40(3), pp. 337-341 https://journals.sagepub.com/doi/10.1177/0361684316654282 .

Corus, C., Saatcioglu, B., Kaufman-Scarborough, C. Blocker, C., Upadhyaya, S. and Appau, S. 2016. Transforming Poverty-Related Policy with Intersectionality in *Journal of Public Policy and Marketing.* vol 35(2) Fall 2016 211-222 https://www.jstor.org/stable/44164853. https://doi.org/10.1509/jppm.15.141

Crenshaw, K. 1989. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics in University of Chicago Legal Forum Volume 1989, Issue 1 Article 8. https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1052&context=uclf

Crenshaw, Kimberley. 1994. "Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color", in The Public Nature of Private Violence, M. Fineman and R. Mykitiuk Eds. Routledge: New York. pp. 93-118 http://www.hsph.harvard.edu/Organisations/healthnet/WoC/feminisms/crenshaw.html

Fischer, A. 2011. *Reconceiving social exclusion.* Brooks World Poverty Institute, The University of Manchester http://hummedia.manchester.ac.uk/institutes/gdi/publications/workingpapers/bwpi/bwpi-wp-14611.pdf

Frontline AIDS. 2019. REAct User Guide https://frontlineaids.org/were-on-the-frontline-of-human-rights/

Hancock, A-M. 2007. When multiplication doesn't equal quick addition: Examining intersectionality as a research paradigm. Perspectives on Politics, 5. bit.ly/45gxH5l. https://doi.org/10.1017/S1537592707070065

Khan, S., Combaz, E. & McAslan Fraser, E. 2015 *Social exclusion: topic guide.* Revised edition. Birmingham, UK: GSDRC, University of Birmingham https://gsdrc.org/wp-content/uploads/2015/08/SocialExclusion.pdf

Krause, H. 2021. *Why We Need Intersectionality in our Demographic Data. Current Events, Data Analysis Concepts.* https://idatassist.com/why-we-need-intersectionality-in-our-demographic-data/

Lorgelly, P., Lorimer, K., Fenwick, E., Briggs, A. 2008. *The Capability Approach: developing an instrument for evaluating public health interventions*. University of Glasgow Section of Public Health and Health Policy https://www.gcph.co.uk/assets/0000/0430/Capabilities_-_full_report__August_08.pdf

Mackay, K. 2007. *How to Build M&E Systems to Support Better Government*. World bank Independent Evaluation Group: Washington. https://www.samea.org.za/documents/Seminar%20documents/How_to_build_ME_gov.pdf. https://doi.org/10.1596/978-0-8213-7191-6

Mbele, T. 2021. *What is meant by transforming evaluations for Africa?* https://twendembele.org/wp-content/uploads/2021/04/MAE-Practice-Brief-Final.pdf

Mena, E., Bolte, G. 2019. The ADVANCE GENDER Study Group. *Intersectionality-based quantitative health research and sex/gender sensitivity: a scoping review*. Int J Equity Health 18, p. 199. https://doi.org/10.1186/s12939-019-1098-8

Morkel, C. 2021. *Made in Africa Evaluation: What is meant by transforming evaluation?* Working Paper. Johannesburg: Twende Mbele.

OECD. 2018. *Development Co-operation Report 2018: Joining Forces to Leave No One Behind*, OECD Publishing: Paris, https://doi.org/10.1787/dcr-2018-en.

OECD and SIGI 2021. *Regional Report for Africa, Social Institutions and Gender Index*. OECD Publishing: Paris, https://doi.org/10.1787/a6d95d90-en.

Ontario Human Rights Commission. 1 July, 2023. *Applying an intersectional approach* http://www.ohrc.on.ca/en/intersectional-approach-discrimination-addressing-multiple-grounds-human-rights-claims/applying-intersectional-approach

Palmen, R. 2021. *Act on Intersectionality*. Zenodo. http://doi.org/10.5281/zenodo.5602859

Riddell, RC. 2014. *Does foreign aid really work? An updated assessment*. Development Policy Centre Discussion Paper 33, Crawford School of Public Policy, The Australian National University, Canberra. https://doi.org/10.2139/ssrn.2409847

Rohwerder, B. 2016. *Poverty and Inequality: Topic guide*. Birmingham, UK: GSDRC, University of Birmingham.

Rouhani, S. 2014. I*ntersectionality-informed Quantitative Research: A Primer*. Institute for Intersectionality Research and Policy.

Runyan, A. 2018. What Is Intersectionality and Why Is It Important? Building solidarity in the fight for social justice. *American Association of University Professors* November-December 2018. https://www.aaup.org/article/what-intersectionality-and-why-it-important#.Yb9PdmhBzrc

Shanker, V. 2018. *What is intersectionality and how can evaluators do it?* American Evaluation Association. https://aea365.org/blog/fie-tig-week-what-is-intersectionality-and-how-can-evaluators-do-it-part-1-situation-analysis-by-vidhya-shanker/

Shulman, K. Harless, J. Kerner, J. Sistrunk, S., Gersh, B., Dube, R., Taleghani, C., Burke, J., Williams, S., Eisenberg, J., Escarce., J. 1999. The Effect of Race and Sex on Physicians' Recommendations for Cardiac Catheterization. *The New England Journal of Medicine*. 340, pp. 618-626. https://www.nejm.org/doi/full/10.1056/nejm199902253400806

Statistics South Africa. 2014. *The South African MPI: creating a multidimensional poverty index using census data*. Statistics South Africa: Pretoria. http://www.statssa.gov.za/publications/Report-03-10-08/Report-03-10-082014.pdf

Symington, A. 2004. *Intersectionality: A Tool for Gender and Economic Justice, Facts and Issues*. The Association for Women's Rights in Development (AWID).

The Global Fund. 2019. *Advisory Review Removing human rights-related barriers: Operationalizing the human rights aspects of Global Fund Strategic Objective 3*. OIG. Geneva. https://www.theglobalfund.org/media/9020/oig_gf-oig-19-023_report_en.pdf

The Opportunity Agenda. 2017. *Ten Tips for Putting Intersectionality into Practice.* https://www.opportunityagenda.org/explore/resources-publications/ten-tips-putting-intersectionality-practice

Chapter 2

Chapter 4: Monitoring Budgets for Accountability and Outcomes

Pundy Pillay

Historically, in both industrialised and developing countries, Ministries of Finance and their implementing departments were primarily concerned with ensuring that spending departments at the national and sub-national (e.g., provinces, municipalities) levels did not exceed their budgets (and, to a lesser extent, that they did not under-spend). Little attention was paid to the expenditure outcomes and whether spending agencies achieved their policy objectives.

Fortunately, this situation is changing. Increasingly, across both developing and industrialised countries, greater attention is being paid to measuring the outcomes of government expenditure to get a better sense of whether government policy objectives are being achieved.

To achieve the objective of measuring performance and outcomes, the budgeting process itself has undergone a fundamental transformation to measure outputs and outcomes to provide a better sense of what is being achieved in different sectors concerning expenditure performance. This has resulted in a consistent movement away from the focus being solely on budgetary inputs and an increasing focus on outputs and outcomes. Such a transformation in the budgeting process has been essential to inform policymakers about the relationship between budgetary expenditure and outcomes. In this way, policymakers will better understand the extent to which national, departmental, provincial, and local objectives are being attained through the budgeting process. Thus, the transformation of the budgeting process to increasingly reflect the expenditure outcomes has demanded greater accountability from policy implementers regarding their spending policies.

This chapter explores the South African case study of budgeting, accountability, and outcomes. Reference is also made to progress in this regard in OECD countries and in another developing country, namely, Mauritius, where the Performance and Programme-based budgeting system is well-advanced.

The Auditor-General of South Africa, Tsakani Maluleke, drew attention to the relationship between spending and service delivery (Sunday Times 2021). Ms Maluleke

made the following pertinent points about the relationship between expenditure and performance:

> "Even though in many instances poor audit outcomes are correlated with poor service delivery, this is not always the case. And herein lies the conundrum. An example is the Overstrand District Municipality, which has for at least three consecutive years received the best audit possible, unqualified with no findings. However, when the lived reality of people in the municipality is observed, service delivery is uneven and in some instances poor. As South Africa prepared for the local government elections, instances such as these received much attention. There is understandably a lot of emotion (reflected in service delivery protests) attached to a lack of services, because service delivery is linked to human rights. Failure to deliver services denies people their human rights and the chance to improve their lives. That public money is provided to government departments and entities to deliver services, to restore the dignity and human rights of millions, attests to the commitment of the government to deliver on this mandate. However, without strong leadership and processes, monitoring and evaluation of the money will not be efficient and people will suffer. A clean audit indicates that the basic aspects of accountability and transparency have been met. This determination is based on whether an auditee is able to show what it has done with its funds, how it has performed against the criteria approved by the legislature and whether it has operated within the rules and laws relevant to the management of public funds. A clean audit is an important milestone in the journey towards building an institution that is effective in delivering on its mandate and can be trusted to continue to do so."

As the Auditor-General points out, a clean audit does not necessarily confirm good service delivery. However, it indicates a solid foundation for service delivery that can benefit the people. The debates about audit outcomes, particularly when these do not seem to correlate with the lived experiences of citizens, have led to an increased focus on bringing the two closer to each other.

Types of Budgeting

In this section, the various types of budgets are described to demonstrate the extent to which there are linkages between budgeting, outcomes, and accountability.

Budgeting in the public sector is complex for a variety of reasons. First, the budgeting process has to collect information from various sources. In a politically decentralised system such as South Africa, budgetary data must be gathered from three government spheres: national, provincial, and local. In addition, other

important actors are interested in the budget, such as the private business sector, trade unions, and other segments of civil society. The budgeting process, therefore, must account for information from multiple sources, diverse interest groups, and different perspectives from within and outside government.

Within the government, there is often tension around the budgeting process between the centre (Ministry of Finance/National Treasury) and line departments. In South Africa, this process may be further complicated by the presence of other coordinating departments in government, such as the Department of Planning, Monitoring and Evaluation (DPME) in the Presidency, and the Department of Cooperative Governance and Traditional Affairs (CoGTA). Furthermore, there are inherent tensions between planners and financial managers within line departments.

The budgeting process is also challenged by allocation issues between different types of spending, for example, current versus capital expenditure.

Unlike the private sector, public sector budgeting is characterised by the absence of a "bottom line", i.e., the need to make profits. Moreover, funding sources can be limited, especially if economic growth is low or slow or the government is reluctant, for whatever reason, to raise taxes. In allocating funds to various departments, there is the added challenge and tension between addressing questions of efficiency versus equity in resource allocation. Finally, there is the issue of political choices versus optimal policy outcomes.

Incremental Budgeting

Given the challenges described above, there is a tendency to resort to "incremental line-item" (ICL) budgeting. Why is ICL budgeting adopted? Budgeting is often very complex, with many interrelated items and competing options. However, ICL budgeting concentrates on changes in various input items, such as personnel and equipment rather than programmes. It involves a narrow range of increases or decreases. For this reason, countries may prefer what they consider a "simpler" system.

In this form of budgeting, decision-making is reduced to concentrating on changes in various input items, such as personnel, equipment, maintenance, utilities, rather than looking at programmes. However, the challenges are many and include departments' overstated funding requests. Where Departments of Finance are not strong, ceilings may not be rigid, and thus may not be taken seriously.

Developing countries sometimes have separate development (capital) and recurrent budgets, often aligned to the type of donor funding, which can become complicated over time. ICL can cause budgets to become unsustainable over time, especially when inflation is high. In developing countries, the scarcity of resources suggests that budgeting must be closely aligned with priorities.

"Traditional Budgeting" (TB) is very similar to ICL. It is a method of preparation of the budget in which last year's budget is taken as the base. Only those items in traditional budgets that are over and above the previous year's budget need to be justified. This process has some advantages: easy to implement; brings stability to the functionality of the organisation; allows consolidation of projects into one single larger one; and is easy to prepare.

However, there are some disadvantages as well: fixed, rigid budget; less motivation as budgets are prepared by top management (bureaucratic); excessive reliance on past year's data; deliberate increase in budgeting costs; and no priority for allocation of resources.

In summary, incremental line-item or traditional budgeting is less concerned with outputs and outcomes. This suggests, therefore, that there is a weak linkage between budgeting and accountability.

Zero-Based Budgeting

Zero-based budgeting (ZBB) is a process that allocates funding based on programme efficiency and necessity rather than budget history. Budgets are not connected to prior year spending; instead, budgets are tied to specific activities and service levels, and as such, spending increases or cuts are not simply spread evenly across budgets. Funding is targeted more to activities that align with the strategy. ZBB is a method of budgeting in which all expenses must be justified for each new period.

As the term suggests, zero-based budgeting starts from a zero-base and every function within an organisation is analysed for its needs and costs. In the end, evidence suggests that zero-based budgeting is challenging to implement. Also, lower-priority programmes do not necessarily receive lower funding.

Moreover, there is no link between budgeting, outputs, outcomes, and accountability in this form of budgeting.

Programme Budgeting

Programme budgeting, developed initially in the United States, is the budgeting system that, contrary to conventional budgeting, describes and gives the exact costs of every activity or programme to be carried out with a given budget. For example, expected results in a proposed programme are described thoroughly, along with its necessary resources, raw materials, equipment, and personnel costs. The sum of all activities constitutes the Programme Budget, and it is relatively easy to find out what precisely will be carried out, at what cost, and what the expected results are.

The formulation phase of programme budgeting involves the preparation of budget proposals by line ministries in negotiation with the central budget authority. This leads to an appropriations bill detailing annual expenditures to be approved by the

legislature. How expenditures are presented for approval and eventually executed varies depending on the type of classification.

Across the OECD, there has been a move towards the classification around programmes, which group expenditures with related policy objectives. Compared to an input-based classification, which specifies the inputs required to provide public goods and services, programme budgeting directs resource allocations towards the results of public spending (OECD 2019).

The inputs and programme budget classifications in the health context, for example, can be described as follows:

Inputs: Wages; goods and services; transfers; capital expenses.

Programmes: Health protection and promotion; maternal health; improved hospital service; ministry support services.

The advantages of a programme-based budgeting process such as this are as follows:

- Health ministries can actively engage in the definition of programmes. This shifts the focus away from inputs required to provide health services towards the objectives. Moreover, by engaging health officials, budgetary decisions will more closely align with health sector priorities.

- Rather than rigid output controls, managers have greater flexibility over programme funds so that spending can be redirected as health needs change. Greater control over the choice of inputs for health officials can also increase the efficiency of public spending.

- Programmes provide a framework for accountability by holding programme managers accountable for results. Programmes can also increase the transparency of how public funds are spent (OECD 2019).

However, several factors must be considered for the successful implementation of programme budgeting, for example:

- How should programme budgets be designed?

   OECD countries use a hybrid system to design programmes. In the health sector, programmes are designed around health policy objectives, the type of service, or administrative and support services. In general, disease-specific programmes or interventions are integrated into broader programmes. Ultimately, the design should reflect the priorities and responsibilities of the health ministry or entity, creating a clear link between the funds and the programme activities and objectives.

- What are the risks?

   Programme budgeting reduces the control over inputs for the Ministry of Finance, increasing risk by allowing opportunities to misuse budget funds. To mitigate this concern, some countries that have moved towards programme budgeting have

maintained separate line-item controls for certain expenditure items, such as the administrative costs of core ministries.

- What are the prerequisites?

    A programme budget does not eliminate the need for other types of classification, and information on inputs must be available to cost programmes fully. Without this, comparing programme alternatives becomes problematic. Programme budgeting also requires strong cooperation of ministries of health, for example, to define the scope of programmes (OECD 2019).

In summary, programme budgeting (PB) represents a major step away from focusing only on inputs to increasing emphasis on programmes and outputs. PB represents an important first step in the evolution of the budgeting process towards greater accountability. This accountability is further strengthened by the development of the performance-based budgeting system.

Performance Budgeting

As Shah and Shen (2007:37) put it,

> "...starting in the 1990s, first in industrialised countries and later in several developing countries, performance management and budgeting reforms have been undertaken to transform public budgeting systems from control of inputs to a focus on outputs or outcomes, in the interest of improving operational efficiency and promoting results-oriented accountability."

Performance-based Budgeting (PBB) is the practice of developing budgets based on the relationship between programme funding levels and expected results from that programme. The PBB process is a tool that programme administrators can use to manage more cost-efficient and effective budgeting outlays.

PBB focuses on the "results" and asks "why the money is spent". Its advantages include: (a) set accountability; (b) clear purpose; (c) improvement in performance; and (d) transparency. Its disadvantages include: (a) it is subjective; (b) it needs a robust system of evaluation; (c) manipulation of data is possible; and (d) it can be difficult for long-term projects.

Box 1 provides an example of performance budgeting in Australia by examining the child-care support programme of the country. The programme shows clearly-defined performance objectives, sub-programmes, and policy measures – the effectiveness of targeting; measures of quality – access and choice; quality assurance; quantity of services provided; and cost.

Box 1: Performance Budgeting: Australia's Child Care Support Programme

Performance Objectives

- Promote, support, and enhance quality child-are.

- Improve access to childcare for children and families with special or additional needs.

- Support equitable access to childcare for children and families in areas or circumstances where services would not otherwise be available.

Australia's Child Care Support Programme includes the following sub-programmes:

- Child Care Benefit

- Jobs Education and Training Child Care Fee Assistance

- Stronger Families and Communities Strategy – Choice and Flexibility in Child Care

- Support for Child Care

The fourth sub-programme above, Support for Child Care, is funded by payments made directly to providers and the states. This programme was introduced in 1997 to encompass all the ongoing and new programmes the department funds to support childcare.

Table 4.1:   Shah and Shen, 2007:143

Measure	Number or percent
Effectiveness: Targeting	
No. of children with additional needs using Australian govt-approved childcare services	126 000
Children with disabilities	16 700
Indigenous children	15 000
Children from non-English backgrounds	95 000
Quality: access and choice	
No. of children with disabilities assisted into mainstream activities	16 700
Quality assurance	
Percentage of centres satisfactorily participating in the Quality Improvement and Accreditation System	90
Percentage of family day care services satisfactorily participating in Family Day Care Quality Assurance	90
Quantity	
Number of Indigenous services	270
Number of services in rural areas	1 200
Price:	$A 18.4 million

Performance monitoring frameworks are often associated with the introduction of programme budgeting. Indicators track the performance of budget programmes' pre-defined outcomes and strengthen accountability for results. This link between budget programmes and key performance indicators provides a framework for integrating performance and monitoring into the budget. Many analysts now refer to Performance and Programme Based Budgeting (PPBB) for these reasons.

PPBB relates to obtaining measurable results and is more effective for better service delivery and improving value for money in public spending, thus making governments more accountable. An increasing number of African countries are adopting PPBB. The outputs are goods and services that a department or spending agency delivers to its citizens, and the outcomes are the intended impacts of outputs.

Inputs refer to labour, capital goods, and raw materials required to produce outputs. The efficiency of government spending can be measured by the "output-input ratio". The effectiveness of government spending, on the other hand, refers to the extent to which inputs, outputs, and outcomes achieve objectives or impact.

In summary, PPBB is ideal for measuring the linkages between spending, expenditure outcomes, and accountability. PPBB ensures greater accountability on expenditure to both legislatures and taxpayers, and it results in improved resource allocation and management as well as enhanced efficiency in public service delivery. In essence, PPBB moves public budgeting from the control of inputs to focus on outputs and outcomes. PPBB has a two-fold rationale: improving operational efficiency and promoting results-oriented accountability.

Performance budgeting presents the purpose and objectives for which funds are required, the costs of programmes, and the outputs to be produced under each programme. PPBB comprises a results-based chain, such as the following:

- Inputs and intermediate inputs – resources to produce outputs.
- Outputs – quantity and quality of goods and services produced.
- Outcome – progress in achieving program objectives.
- Impact – programme goals.
- Reach – people who benefit or hurt by a programme.

The case of Performance-Based Programme Budgeting can be more clearly illustrated using the example of education (Shah and Shen 2007:144-145):

- Programme objectives – improve quantity, quality, and access to education services.
- Inputs: educational spending by age, gender; urban/rural; spending by grade level; and number of teachers, staff facilities, tools, books.
- Intermediate inputs: enrolment, student-teacher ratio; class size.

- Outputs – achievement scores; pass/graduation rates, drop-out rates.
- Outcomes – literacy rates, supply of skilled people.
- Impact: informed citizenry; civic engagement; enhanced international competitiveness.
- Reach: winners and losers from government programmes.

To reiterate, the main benefits of PPBB are the following: higher transparency and accountability, more informed budgetary decision-making, improved management in government agencies, and better communication between budget actors and citizenry.

The Medium-Term Expenditure Framework (MTEF)

The Medium-Term Expenditure Framework (MTEF) enables a government to prioritise spending effectively. In South Africa, the MTEF provides budgetary amounts for spending agencies (e.g., national departments, provinces, and municipalities) for three years. The great advantage of an MTEF is that it enables better-planned expenditure (particularly capital expenditure) because spending agencies are provided with indicative spending allocations for three years rather than for one year.

The National Development Plan (NDP) and the Medium-Term Strategic Framework (MTSF)

In South Africa, the long-term development plan is embodied in the National Development Plan (NDP) (NPC 2012), which sets out the development priorities for the country until 2030. The Medium-Term Strategic Framework (MTSF) sets out the government's policy priorities for a five-year period aligned to the electoral cycle. The MTSF is developed by all government spending agencies (national departments and provinces). Provincial MTSFs should be informed by the Integrated Development Plans (IDPs) of municipalities in each province.

Since 2010, the government's planning frameworks (first the MTSF and then the NDP) have increasingly influenced the budgeting process. In other words, the government sets out its priorities which are then funded through the MTEF and the annual budgeting process.

Performance Budgeting in OECD Countries

This section draws from the OECD (2019) publication, *OECD Good Practices for Performance Budgeting*, which provides a comprehensive review of the budgeting processes in OECD countries while emphasising "good practices".

Performance budgeting is defined by the OECD (2019:9) as:

> "...the systematic use of performance information to inform budget decisions, either as a direct input to budget allocation decisions or as contextual information to inform budget planning, and to instil greater transparency and accountability throughout the budget process, by providing information to legislators and the public on the purposes of spending and the results achieved."

Several "Good Practices" described by the OECD document are summarised below.

Good Practice 1: the rationale and objectives of PB are clearly documented and reflect the interests of key stakeholders (2019:23-27).

- The rationale, objectives, and approach to PB are set out in a strategic document such as an organic budget law of PFM reform programme (e.g., New Zealand).
- The interests and priorities of multiple stakeholders in the budget cycle are reflected in the objectives and design of the PB system (e.g., France, Australia, the United Kingdom).
- PB is championed by political leaders, with support from senior officials (e.g., Canada).
- The introduction of PB is supported by regulations and guidelines (e.g., Australia).

Good Practice 2: PB aligns expenditure with the strategic goals and priorities of the government (2019:29-32).

- Budget proposals are systematically linked to relevant development plans, government programme commitments, and other strategic direction and priority statements.
- Multi-year budget frameworks provide realistic and reliable fiscal parameters for preparing performance budgets.
- Central government activities and budgets support the achievement of complex objectives requiring inter-ministerial collaboration.

In this regard, there are essential factors to note with relevance to South Africa and other developing countries. One of these is the link to strategic plans providing alignment between budget and the government's policy priorities. Another factor to consider is that MTEFs provide "structured approach to integrating fiscal policy and budgeting over a multi-year horizon that links fiscal forecasting, fiscal objectives or rules and planning of multi-year budget estimates. MTEFs can improve the effectiveness of public spending by aligning public expenditure with national priorities and giving government agencies greater certainty of resource availability over multi-year periods, promoting more effective forward planning and resourcing of policies that require an extended time horizon for implementation, such as large capital projects" (OECD 2019).

Good Practice 3: The PB system incorporates flexibility to handle the varied nature of government activities and the complex relationships between spending and outcomes (2019:33).

- The type and volume of performance information required varies based on the nature of the programme.

- Government uses a mix of performance measures reflecting the multi-dimensional nature of performance in the public sector.

- Programme structures are aligned with the administrative responsibilities and service delivery functions of ministries and agencies.

- Expenditure classification and control frameworks are revised to facilitate programme management and promote accountability for results.

Key to achieving the above is choosing the right mix of performance indicators. The OECD defines good indicators as:

- Limited to a small number for each policy programme or area.

- Clear and easily understood.

- Allow for tracking of results against targets and comparison with international and other benchmarks.

- Make clear the link with government-wide strategic objectives.

A common challenge facing many OECD countries is the identification of a balanced set of indicators that reflect the multi-dimensional character of performance in the public sector (2019:33). Key dimensions of performance that need to be considered in the OECD view are the achievement of key government policy goals, delivery of high-quality public services, value for money, and compliance with internal business rules. France and Australia are good examples in this regard.

Good Practice 4: Government invests in human resources, data, and other infrastructure to support PB (2019:41).

- The Central Budget Authority (CBA, usually the Ministry of Finance) builds capacity, internally and within line ministries, to manage and operate the PB system.

- The CBA regularly reviews and adjusts the operation of the PB system to improve its performance.

- Performance measurement systems are progressively improved to provide quality data on a reliable basis.

- Performance data is governed and managed as a strategic asset to ensure that the data is discoverable, interoperable, standardised, and accessible timeously.

Good Practice 5: PB facilitates systematic oversight by the legislature and civil society, reinforcing government performance orientation and accountability (2019:47).

- Annual budget and expenditure reports presented to the legislature contain information about performance targets and levels of achievement.

- The supreme audit institutions (SAIs) carry out performance audits, including tests of the accuracy and reliability of reported performance.

- Parliament, supported by the SAI, scrutinises performance-based budgets and financial reports, holding ministers and senior public managers accountable in the event of poor performance or misrepresentation.

- Accessible formats such as online performance portals and citizen budgets help citizens, civil society, and the media to monitor performance.

The United Kingdom, Canada, and Mexico are good examples of the above.

Good Practice 6: Performance budgeting complements other tools designed to improve a performance orientation, including programme evaluation and spending reviews (2019:49).

- Ex ante appraisal of new spending programmes is used to strengthen programme design, including key performance indicators, and to facilitate processes of monitoring and ex post evaluation (e.g., Chile).

- Ex post evaluations of major spending programmes are carried out on a rolling basis and the findings are systematically fed back into the budget preparation process.

- Spending reviews are used in conjunction with PB to review the justification for spending and to identify budgetary savings that can be redirected to support priority goals. A good example of linking performance measurement and evaluation is Canada.

Good Practice 7: Incentives around the performance budgeting system encourage performance-oriented behaviour and learning (2019:53).

- The centre of government promotes a management culture that focuses on performance.

- Performance management comparison and competition between similar entities as a means of improving effectiveness and efficiency in service provision must be encouraged.

- Identified individuals and teams are responsible and accountable for the achievement of performance goals.

- Managers organise structured internal discussions to review financial and operational performance regularly through the year.

- Responses to programme under-performance emphasise learning and problem solving, rather than individual financial rewards and penalties.

Programme–Based Budgeting in Mauritius

According to CABRI (2013a:1), many African countries follow the international trend of "introducing a performance orientation into annual budget processes and planning budget expenditures over a multi–year period". A major aim in this regard is to closely align strategic socioeconomic planning with annual budgets and medium–term budgetary frameworks (MTBFs). In francophone Africa, programme budgeting has been chosen as the preferred 'model' to performance budgeting. Elsewhere in Africa, although programme-based budgeting is being introduced in some countries, there is greater diversity in the emerging performance-based budget systems (CABRI 2013).

As stated by CABRI (2013), over eighty percent of African countries were introducing, or were committed to introducing, some form of PPBB. In many cases, PPBB reforms were adopted as part of a broader package of public financial management (PFM) reforms. The decision to introduce PPBB often originated in each country because of pressure from regional bodies in Africa or the donor community. PPBB reforms were introduced mainly due to the results-oriented approach to budgeting being perceived as a major means of improving expenditure reallocation, particularly towards social sectors that contribute to poverty reduction.

Mauritius has one of the most advanced forms of budgeting in the developing world. The focus here is on performance in the public sector. A PPBB budgeting system has been adopted, in which the budget management process links resources to clear and agreed outcomes and outputs to ensure greater accountability in terms of performance.

The rationale for the adoption of PPBB involved the assurance of greater efficiency of public expenditure, as well as greater emphasis on transparency and performance. Other intended outcomes included the linking of PBB to a three-year MTEF, focus on programme outcomes and performance, and a shifting of resources to performance areas (CABRI 2013b).

Considerable effort was made in raising awareness of PBB's potential as an effective planning tool to achieve the goals of ministries and departments. In Mauritius, PBB has been built based on an efficient planning and budgeting system. Moreover, the emphasis in Mauritius has been on simplicity and effective functioning. A standard and short format was adopted for the PBB across all ministries, and a simple programmatic classification with an initial focus on outcome indicators was introduced. Mauritius also initiated parallel reforms through a public sector investment programme, procurement, and the development of appropriate human resources.

Three main areas were the focus in the development of the PBB system, these being strategic direction, performance information, and accountability. There was

Chapter 4

a strong strategic orientation in the budgeting process, with better links between planning and budgeting. A results-oriented public financial management system was developed for improved efficiency and effectiveness in resource utilisation. Better transparency and accountability were initiated on the part of ministries and departments who became more involved in budgeting and planning. The role of the National Audit Office, which audits performance information, was also stressed.

While Mauritius appears to have addressed many of the implementation challenges relating to performance-based budgeting, Robinson and Last (2009) draw attention to some of the challenges that continue to plague many other developing countries.

The first point that Robinson and Last (2009:8) make in this regard relates to the failure to view PBB as part of a broader set of reforms, which they consider to be a necessary condition. The reforms they refer to include reforms of the civil service to "increase motivation and incentives of public employees", undertaking "organisational restructuring to increase the focus on service delivery and improve coordination (e.g., creation of agencies and reduction of the number of ministries)", and introducing "institutional and oversight changes to strengthen public accountability for performance".

Linking Budgets to Performance and Accountability – South Africa

Together with Mauritius, South Africa has likely advanced the furthest amongst African countries in the development and implementation of Performance and Programme-based Budgeting. However, there is still a long way to go in South Africa, particularly with respect to measuring performance in relation to spending.

In the Department of Basic Education, for instance, serious efforts are being made in the development of indicators for monitoring attendance at early childhood development facilities, quality of education, grade repetition, school attendance, pass rates, and implementation of the national school feeding programme (DBE 2021). However, there is very little information and data linking these measures to actual spending, so the link between spending, outputs, and performance outcomes is still largely absent.

Similarly, the Department of Health has developed several indicators to monitor performance, which include mortality rates, immunisation coverage, and incidence of HIV/AIDS and tuberculosis (DoH 2021). As described above in the case of education, the link between output, outcomes, and expenditure has not been developed to any significant extent.

One of the major challenges to the development of appropriate performance measures is the highly decentralised political system that has been developed in South Africa in the post-apartheid era comprising national government, provinces, and municipalities.

A key institutional mechanism for monitoring expenditure and outcomes, inter alia, was created through the establishment of the Department of Planning, Monitoring, and Evaluation (DPME) in the Office of the President in 2010. For the first time in the democratic era, an opportunity arose to implement appropriate structures and mechanisms for more effective measurement of the relationship between expenditure and outcomes in all three spheres, namely, national, provincial, and local. In addition to the national DPME, there are also provincial equivalents established in the Premier's Office providing a strong institutional link in this context between the national and provincial governments.

However, insufficient progress has been made in fully developing the linkages between expenditure, outcomes, and performance. A fundamental reason for this, alluded to earlier, is the constitutional structure of national government, provinces, and municipalities. While the constitution provides for cooperation between the three spheres through the principle of 'cooperative governance', it also ensures a significant degree of autonomy for provinces and municipalities. This is particularly true in the way in which spending priorities are determined, and resources are allocated.

In the education and health sectors, for example, policy is made at the national government level but implemented by the provincial departments whose political heads report to the provincial Premier and not to the national Ministers of Education and Health respectively. In a similar vein, the financial resources for the implementation of education and health policies are determined through the provincial budgeting process. Thus, for those functions that are shared between national and provincial governments (such as education and health), the onus for measuring performance outcomes should be on the implementing agencies (in this case, provinces). However, in most cases, provinces appear to be unprepared, unwilling and/or incapable of undertaking effective measurement of budgeting outcomes at this stage.

The challenge of measuring performance outcomes is even greater at the local government level where the human and institutional capacity to implement policy is severely lacking, making the challenge of measuring outputs and outcomes an even greater one than it is at the national and provincial levels. The inefficiency in this regard has been noted with great eloquence by South Africa's Auditor General. Govender and Reddy (2019) in their case study, cite inappropriately defined projects and programmes as well as unreliable data as major barriers to the development of an effective performance-based budgeting system in the eThekwini metropolitan municipality in KwaZulu-Natal.

The national government also provides significant funding to both provinces and municipalities in the form of conditional grants. For provinces, funding is provided in this form, inter alia, to the education, health, and agricultural sectors. In local government, large conditional grants are provided for infrastructure. Nevertheless, the measurement of the "performance" of these grants is also a matter of concern given the importance of infrastructure, health, and education, inter alia, for South African development.

In conclusion, while South Africa has made significant progress towards reforming its budgeting processes to reflect a better relationship between inputs, outputs, and outcomes through its adoption of Performance and Programme-Based Budgeting, there is still some way to go to in terms of convincing taxpayers and citizens in general that government is making the best possible use of the financial resources at its disposal. Poor service delivery outcomes at all three spheres of government might suggest otherwise.

As Maluleke (2021:18) put it: "A clean audit is not necessarily a confirmation of good service delivery, but it indicates a solid foundation for service delivery that can benefit the people. The debates about audit outcomes, particularly when these do not seem to correlate with the lived experiences of citizens, have led to increased focus on bringing the two closer to each other".

Conclusion

As noted in the introduction of this chapter, Ministries of Finance and their implementing departments were largely concerned with ensuring that spending at the national and sub-national levels did not exceed their budgets, and that they did not under-spend. Little attention was paid to the outcomes of expenditure and whether spending agencies were achieving their policy objectives.

As has been discussed, a change in this situation can be seen. Greater attention is being paid to measuring the outcomes of government expenditure to get a better sense of whether government policy objectives are being achieved.

This chapter discussed the nature of the complexities associated with public sector budgeting and the tensions between Ministries of Finance and line departments. Various types of budgeting were evaluated, including incremental line-item, traditional, zero-based, programme-based, and performance-based budgeting. The links between the programme- and performance-based approaches and related issues such as the MTEF, the NDP, and the MTSF relating to South Africa were highlighted.

The experience with performance-based budgeting in OECD countries was examined in some depth with seven sets of good practices described. The successful experience with PPBB in Mauritius was then reflected on, given that this country has one of Africa's most advanced budgeting systems. Section 5 noted the progress made by South Africa in linking budgeting to performance and accountability while underlining some of the ongoing challenges in important sectors such as education and health, and the insufficient progress that has been made in developing the linkages between expenditure outcomes and performance. The complex intergovernmental system comprising national, provincial, and municipal governments and the associated lack of human and institutional capacity are linked to many of these challenges.

Chapter 4

Reference List

Auditor-General of South Africa. 2021. Integrated Annual Report 2020-21 – *Increased Relevance in a Changing World*: Pretoria.

Collaborative African Budget Reform Initiative. 2013a. *Performance and Programme-based Budgeting in Africa*: A Status Report. Pretoria

CABRI. 2013b. *Programme-Based Budgetary Reform in Mauritius.* Briefing Paper no.7

Department of Basic Education. 2021a. *Annual Performance Plan 2021/22.* Pretoria. March

Department of Basic Education. 2021b. *Annual Report 2021/22.* Pretoria. March

Department of Health. 2021. *Annual Performance Plan 2020/21.* Pretoria. March

Department of Planning, Monitoring and Evaluation .2021. *National Evaluation Plan 2020-2025.* Pretoria. March

Govender, N. and Reddy, P.S. 2019. *Performance Budgeting in Local Government – a case study of eThekwini Municipality in South Africa, in Performance-based Budgeting in the Public Sector.* Eds. De Vries, M.S., Nemec, J. and Spacek, D. Palgrave-Macmillan, Governance and Public Management series. https://doi.org/10.1007/978-3-030-02077-4_7

Maluleke, T. 2021. *Keeping an eye on the public purse.* Sunday Times.

National Planning Commission. 2011. *The National Development Plan*, The Presidency, Pretoria

OECD. 2019. *OECD Good Practices for Performance Budgeting.* OECD Publishing: Paris

Robinson, M. and Last, D. 2009. *A Basic Model of Performance-Based Budgeting.* International Monetary Fund: Washington, D.C. https://doi.org/10.5089/9781462320271.005

Shah, A. and Shen, C. 2007. *A Primer on Performance Budgeting in Budgeting and Budgetary Institutions.* Ed. Shah, A. Public Sector Governance and Accountability Series, The World Bank:Washington, D.C. https://doi.org/10.1596/978-0-8213-6939-5

Section 2

INSTITUTIONALISING MONITORING SYSTEMS

Chapter 5: Institutionalising Monitoring Systems: Introducing Results-Based Monitoring to Organisational Systems

Khotso Tsotsotso

It has been observed that social development efforts were frustrated by the recent global pandemic in the past three years. The pandemic had a regressive effect both on human and economic conditions globally. It crowded out other initiatives by redirecting public and private resourcing towards managing the effects of COVID-19 (Sheth, Lorig, Stewart, Parodi and Ritter 2021). As a result, with the growing demand for social innovation and solutions that work, the development community has become less tolerant of programmes and investments that struggle to show meaningful impact (Mashamaite 2014; Breakfast, Nomarwayi, and Dodd 2021). These concerns have rekindled interest in good programme management and, more importantly, in robust monitoring systems that can assess how well development projects and social programmes meet their intended objectives over time.

In practice, it is conventionally understood that theories of change lead M&E frameworks, which provide direction to monitoring systems. Contrary to this norm, monitoring is typically a management practice that is part of a system that is often entrenched in institutional patterns, and not necessarily aligned with the learning and evaluation system of an organisation. Although the development of theories of change is arguably easier once management and stakeholder commitment to participate in the process is achieved, embedding a results-based monitoring system within an organisation with existing norms and standards requires significant cultural and behavioural shifts. It involves significant creative and collaborative pushes that may conflict with existing norms, and even experience 'push-back'.

This chapter explores and discusses disparities between organisational norms and behavioural requirements for results-based monitoring systems and their effects on management tensions as organisations attempt to implement and institutionalise

results-based monitoring systems. It discusses lessons from experts and practitioners in South Africa who implemented results-based monitoring systems in different types of organisations.

The chapter first explains monitoring in the context of a programme or organisation to ensure a grounded understanding of results-based monitoring systems. The chapter then explains the concept of systems in the context of an organisation. Finally, to ground the arguments made on lessons from practice, different theoretical perspectives of change in the context of an organisation are explained.

Monitoring in the Context of a Programme or Organisation

Despite the reality that M&E practice can differ depending on the underlying programme or policy context, which can fundamentally alter the content, process, and monitoring activities, practitioners and scholars have some common ground regarding the 'occasionally blurred line' that differentiates Monitoring from Evaluation. Gosling and Edwards offer a helpful description of monitoring as "...a systematic and continuous collection and analysis of information about the progress of a piece of work overtime" (2003:12).

Therefore, monitoring is an intent to keep all stakeholders well-informed and positioned to make real-time decisions throughout the implementation of an intervention. It differs from evaluation in several ways; i) it is traditionally carried out by internal rather than external staff and serves as a management tool, ii) it is performed routinely (ongoing) rather than periodically (at specific intervals), and iii) although not the case in contemporary practice, monitoring tends to focus more on early-stage and process results such as activities and outputs, as opposed to outcomes and impact. However, in response to the call for measuring meaningful change, the contemporary practice of monitoring tends to prioritise tracking outcome indicators (Nxumalo 2016). Due to these three monitoring characteristics, it is usually designed as a system with processes, rules, infrastructure, and people (Simister 2017).

In support of the central argument made in this chapter, Simister (2017) emphasises the need to embed programme monitoring firmly within programme and organisational management norms by clarifying the seven types of monitoring, each monitoring specific aspects of organisational operations. The seven types of monitoring include:

1. **Process and performance monitoring** monitors the inputs, activities, and outputs of the programme or policy. It is intended to ensure that process plans are carried out as stated and to track deviation from plans.

2. **Outcomes or impact monitoring** tracks changes brought about by the programme over time. It monitors progress based on pre-stated outcomes towards the programme objectives.

3. **Beneficiary monitoring or beneficiary contact monitoring** tracks beneficiary sentiments throughout the programme. It is intended to manage motivation to participate and manage attrition risk and can include routine feedback mechanisms and routine surveys.

4. **Situation monitoring.** In results-based systems, the programmes make explicit assumptions about the programme context and situation. The programme might also state environmental risks to the underlying intervention and might monitor ongoing changes in a dynamic context or programme environment based on predefined indicators to manage these risks and assumptions.

5. **Financial Monitoring.** Although this is often managed separately from the programme monitoring system, tracking expenditure and resource levels against progress can be critical in resource-sensitive or pilot programmes earmarked for large scaling.

6. **Administrative or logistics monitoring.** Simister describes this as tracking the maintenance of premises, transport, personnel, stock-keeping, and other forms of administration.

7. **Compliance monitoring.** In many social or environmental initiatives, strict compliance is crucial as it can determine the immediate failure of the initiative. In such programmes, tracking compliance is based on a predetermined set of indicators.
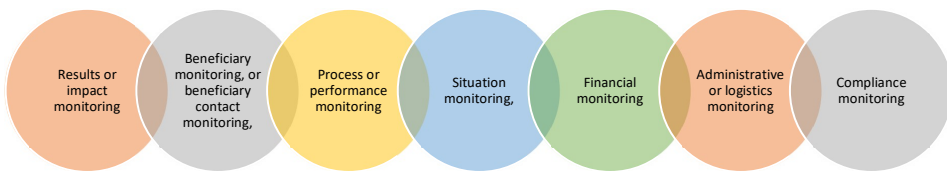
Figure 5.1:   Simister, The Seven Forms of Monitoring, 2017

Systems in the Context of an Organisation

In the complex fourth industrial revolution era, the word "system" is used interchangeably to refer to a digital platform that can be seen. Otherwise phrased, the term is used to describe the conceptual interconnectivity of processes, people, infrastructure, and rules within an organisation. In this chapter, the interest and focus are on the conceptual meaning of a "system". Therefore, "system" refers to a set of 'things' working together as parts of a mechanism, or an interconnecting network, resulting in a complex whole. Alternatively, a system can refer to a set of principles or procedures according to which something is done; an organised scheme or method (Oxford Dictionary).

McNamara (2017) explains a system aligned to the focus of this chapter. He describes a system as simply an organised collection of parts (or subsystems) that are highly integrated to accomplish an overarching goal. An organisation might consist of a

clear organisational goal that other functions of the organisation (subsystems) collaboratively work towards achieving. These may include administrative and management functions, products and services, and human resource functions. Finally, interactions between these functions determine the culture and behavioural norms of the organisation's system. The key characteristic is the interdependency or the interrelationships between organisational functions. The interactions (based on the known set of rules) between these functions determine the organisation's outputs towards its goals. Similarly, in the context of M&E, the 'monitoring system' typically consists of a stated logic of change (results of change), a clear set of programme results, the indicator framework, data collection tools, the sources of data or information, the users and managers of the information, and the infrastructure. These are intended to interact with each other to ensure continuous learning, accountability, and improved performance.

In this chapter, both the monitoring system and the organisational system should be considered as open systems that must interact with each other and the environment. They continually exchange feedback between each other and their various parts to ensure that they remain closely aligned and focused on achieving the goals of the organisation. The monitoring system needs to interact with broader management processes. These can include knowledge management (KM), financial management, human resources, data storage, marketing, fundraising, or quality assurance. For example, individuals rolling out a new monitoring system may need to rely on KM and data storage processes to ensure that the system operates appropriately (Feruglio and Nisbett 2018). Alternatively, and more critically, monitoring information may need to inform management decisions, thereby adding a step to decision-making processes.

Now, suppose the parts of either the system or the systems' activities are altered or forced to deviate from the norm; as a reaction to the change, the system should be expected to make necessary adjustments to achieve the goals for which it was designed.

This chapter argues that altering any of the rules, parts, and system resourcing will inevitably also either weaken or strengthen the system's functioning, at least in the short to medium term (Teece 2018), thereby affecting the nature and quality of the regular system outputs. In practical terms, introducing results-based monitoring processes (the monitoring system) into an existing organisational or programme context with its norms and culture can alter ongoing activities and outputs in the short to medium term. Furthermore, unless the process is managed well, the altering effect can, in turn, result in 'push back' or resistance from management and staff. This phenomenon will be discussed later in the chapter.

Change in the Context of Organisations

M&E practitioners and scholars have provided examples and cases that eloquently describe processes and common agreements about critical milestones for developing results-based monitoring systems, including those described by Kusek and Rist (2001), Clements (2005), Mackay (2006), Kusek (2010), as well as Mapitsa and Chirau (2019). However, there is little literature and few theories detailing and modelling organisational change tensions brought about by the introduction of results-based monitoring systems. Although there is a significant history of literature and a plethora of theories that unpack and model general organisational change, the M&E community has not yet sufficiently applied or built from these theories to foster understanding regarding tensions that arise from the implementation of results-based monitoring.

To respond to this 'gap' and ground the chapter's central argument, this section explains key organisational change concepts that help explain tensions experienced during the implementation of result-based monitoring systems.

In this section, organisational change refers to the transformation of an organisation between two points in time (Barnett and Carroll 1995). The trigger for such transformation is usually intentional and explicit, such as the introduction of new technology, new production methodology, and new products, services, or interventions. Or, more important to the topic of this chapter, an introduction of a results-based monitoring system. Throughout the section, change is to be considered in two dimensions; the first dimension is the content of change, which is the consideration of the 'what' of change, i.e., what elements of the organisation (the elements of structure or radical shifts in a single element of structure) alter during the transformation. The second element of change studies the processes of change. These can be the speed of change, the sequence of change activities, decision-making, communication, and the resistance encountered. It is, therefore, the 'how' of change.

The Perspective of Consequence of Change Theory

One of the principal reasons for the consideration of organisational change theory is the consequence of organisational change. This consideration of organisational change theory serves as a framework for the understanding of successes and failures resulting from the implementation of results-based monitoring systems in an existing organisational or programme context. Understanding this aspect of organisational change theory is particularly critical due to results-based monitoring systems being designed to drive 'core structural change' within any organisational or programme context. Hamman and Freeman (1984) define core structural changes as changes in organisational mission, authority structure, technology, and marketing structure. It should be considered that in practice, any design and development of a results-

based monitoring system starts with review and refinement of an organisation or programme's results, which are typically defined in the form of a theory of change. Result-based monitoring systems often include a retrofitted theory of change with 'smartly' articulated impact and outcome statements in the form of a results chain (McConnell 2019).

This theory-based approach to refining and re-stating the organisation's intended results inevitably alters its mission. One of the key functions and intentions of a results-based monitoring system is to promote and support a data-driven accountability culture. By intention, such a culture mainstreams M&E evidence to drive accountability, management, and strategic decision-making. In other words, a monitoring system aims to change how the organisational authority structures make decisions towards a culture of evidence-based decision making. By defining what should count (the organisation's outcomes) and outlining a clear measurement framework and plan, a results-based monitoring system organically proposes a way in which to change ideal organisation operation. Some monitoring systems might be explicit about intervals of staff reflections, methodology for managing and sharing knowledge, and may define a matching theory of implementation that accommodates results-based monitoring activities (Montague 2019; Jankvist, Gregersen, and Lauridsen 2021; DuBow and Litzler 2019). To substantiate this point, a flagship education programme for the South African Department of Basic Education adopted a result-based monitoring system in 2018 that refined its rhythms for management and accountability engagements, the way measurements are conducted, and the re-alignment of its performance review process. The programme also adjusted its implementation approach to accommodate and integrate real-time data collection during teacher training sessions to feed into the ongoing monitoring cycle, thereby altering the organisational implementation process, and associated reporting and information systems.

One of the benefits and key features of any monitoring system is defining and facilitating the celebration of organisational achievements as defined in the results framework. In fact, with the growth and popularity of results-based management and M&E broadly, the capacity to demonstrate organisational effectiveness and initiative impact through M&E evidence has become a critical requisite for programme or organisational funding. This is especially true for organisations with an overarching M&E system that is a central part of the administrative system, such as development corporations and governments. This said, the use of monitoring evidence for marketing purposes has become common in such organisations.

Therefore, we can sufficiently and safely assert that a results-based monitoring system is designed to change the organisational mission, authority structure, technology, and marketing structure. Consequently, we should align with Hennan and Freeman's theory that "Core Structural Change is precarious and leads to an elevated probability of organisational failure and death" (Hean and Freeman 1984:156)

Hennan and Freeman (1984) argue that there is a positive relationship between core structural changes and the probability of organisational failure. However, the theory makes provision that some organisations can manage such changes well, and in these cases, the significant positive performance can be realised. Perhaps this theory can help explain common observations regarding the introduction of results-based monitoring systems:

- The risk of change failure and organisational death can explain the uneasiness among organisational and programme leadership to fully embrace a result-based monitoring system, which often results in push-back or part-implementation of such a system.

- Reluctance to embed a results-based monitoring system due to the probability of organisational failure can be the reason for the absence of appropriate incentives for M&E and low rates of meaningful learning.

These observations can account for diminishing investments to maintain appropriate resource levels for a functional monitoring system which, in turn, leads to system ineffectiveness over time.

Hage (1999) argues that changes in the environment trigger change choices like organisational form and the movement toward (or away from) an organic form. An organic organisation is defined as an organisational structure where all the employees are placed at equal levels. In this type of flat structure, the interactions and communication are horizontal by nature (Aiken and Hage 1971). Hage (1999) identifies four theories as perspectives for understanding organisational change. These perspectives are introduced below and are used as reference points for describing changes dictated by the introduction of results-based monitoring systems.

The Perspective of Structural Contingency Theory

The original assertion of the structural contingency theory (Burns and Stalker 1961) is that a stable demand leads to a mechanical organisation. In contrast, a changing demand creates the need for an organic organisation, emphasising innovation and flexibility. According to Hage (1999), this assertion was later adjusted by Lawrence and Lorsch (1967) when they argued that an organic form of an organisation is a result of applying new knowledge from the environment and is not necessarily triggered by changes in demand. A more apparent claim of this perspective is that applied knowledge and awareness of shifts in the environment (whether economic or political) should determine needed changes in the organisation that will help match the environment, i.e., reaction to the knowledge of environmental changes determines organisational changes.

The theory explains two observations when applied to our results-based monitoring systems context. The first is that results-based management gained popularity among development corporations, civil society organisations, non-profit organisations,

and governments (Berezovsky, 2017). It increasingly became a standard framework for good management of social initiatives. Due to this view, the donor community made it a condition for fund disbursement, thereby proliferating the practice of RBM in the sector. As a result, a significant shift in the development environment was recognised that warranted a shift at the organisation level (Swiss 2005; Pfeffer and Sutton 2006; Hulme 2007; Kusek 2010; Vähämäki, Schmidt and Molander 2011). In line with Structural Contingency Theory, organisations aware of this shift towards results-based management increased their demand and searched for models that can best respond to this shift in the sector.

The second observation focuses on the magnitude of knowledge and appreciation of its value, as well as application levels. Hage & Powers (1992) argue that more and more knowledge of changes in economic and political sectors determines emphasis on choice and commitment to organisational form. Hage et al. (1993) provide empirical evidence of this relationship. This view provides an important glimpse into the 'maturity spectrum' of adopting results-based monitoring systems between organisations that have introduced them. According to this theory, organisations with superior knowledge of the nature of demand for results-based management and evidence-based decision-making tend to invest in more deepened institutionalisation of results-based monitoring systems than organisations with minimal understanding of the nature of demand from the sector. This emphasises and gives credence to recent and continued investments in RBM and M&E capacity development, especially in the African continent, as a panacea for the poor institutionalisation of results-based monitoring systems. In other words, this view justifies efforts to professionalise M&E as a gateway for improving and deepening evidence use in Africa. In the last two decades, organisations like the Centres for Learning on Evaluation and Results (CLEAR), the Independent Evaluation Group, the African Capacity Building Foundation, Twende Mbele, African Evidence Network, the South African Monitoring and Evaluation Association, African Evaluation Association, and many others have identified and mainstreamed capacity development as a key strategy to improve evidence-based decision-making in the development sectors of the continent.

The Perspective of Political Theory

An implicit assumption of structural contingency theory (SCT) is that managers (as individual staff) will also adjust to meet environmental demands in ways that are appropriate for them. More specifically, SCT relies on the assumption that vital organisational functions or departments that handle critical contingencies of the organisation or programme – referred to by Pfeffer (1981) as "dominant coalitions" – will embrace change and act in ways that accommodate the change. Pfeffer (1981) disagrees with this assumption for two main reasons: Firstly, regardless of what occurs initially (changes in the entire organisation or changes in the 'dominant coalition'),

any shift in the dominant coalition is tantamount to a shift in organisational strategy either for or against the intended organisational change. The dominant coalition can invoke their prerogative to claim potential risk to a core strategy and label the change as high risk, especially when change feels uncomfortable. Following this argument, the true power to organisational change remains mainly with the dominant coalition. For example, in a programme that intends to change teaching pedagogies in schools, the workstream responsible for teacher training is core to such a programme and can therefore be referred to as the dominant coalition. A new theory of change for such a programme might suggest a change in how teacher training is implemented and measured. Now, according to Pfeffer (1981), whether such a change will be successful or sustained by the programme depends directly on whether the teacher training workstream (dominant coalition) accepts it or not. This argument provides an interesting perspective for why many organisations that attempt to implement monitoring systems are met with resistance and hardly see success. Alternatively, in fortunate situations, they show high levels of success.

In practice, it is conventional for drivers of M&E plans and systems (whether internal as officers or external as consultants) to identify 'implementation champions' thoughtfully selected from internal management to act as supporters and leads for change. These may be individuals or groups who have fully embraced the idea of a functional results-based monitoring system. Such champions might not represent the dominant coalition of the organisations as described above. In this case, any proposed activity of the results-based monitoring system that purports to effect fundamental change to the ways of working of the dominant coalition has a high probability to be met with resistance, especially when the change is deemed somewhat uncomfortable. Theory suggests that in this case, the dominant coalition will have the necessary leverage to block the change towards results-based management on behalf of the rest of the organisation. Conversely, where the implementation champion is successfully recruited among influential members of the dominant coalition, fundamental changes to the coalition's ways of working can be received with positive support and thus lead to successful institutionalisation of the results-based monitoring system.

For the second reason, Pfeffer (1974) and Salancik (1978) offer a reminder that the implementation of results-based monitoring systems requires significant resource investment. A compelling point is made which states that while the dominant coalition may enjoy a level of influence and power, they are also beholden to those who control the organisational 'treasure chests'. Real influence rests with those who control organisational resources. Therefore, we are compelled to apply a Resource Dependency lens, which suggests that effective 'championship' for the successful embedding of a results-based monitoring system is best placed with those who control resources. This is typical of development corporations' and governments' procedural standards. A part of the consideration for new or continued funding of

programmes depends on the programme management's ability to demonstrate evidence or a clear intention to implement a results-based monitoring system. As an example, an internal review by Data Innovators (Pty) Ltd of sixteen of its monitoring and evaluation clients in 2020, reflected the fact that the most successful clients (those who showed significant levels of M&E institutionalisation) are organisations in which M&E services are procured and championed by the underlying programme funders (Data Innovators 2020).

The Perspective of Institutional Theory

According to Hage (1999), institutional theory can help unpack how change diffuses within countries and even across them. It can also provide alternative explanations for countries not responding to sectoral or market pressures. Scott (2008) points out that institutions have the responsibility to provide rules and define ways and norms to operate by either discouraging, constraining, or encouraging given behavioural patterns. The central idea is that any organisation operating within the jurisdictions of a defined society will be affected and influenced by society's rules, norms, and behaviour. More importantly, any significant shift in society's rules, norms, and behaviour will, by diffusion, influence the form such an organisation will adopt (Powell and DiMaggio 1991).

Scott (2008) defines the three pillars on which societal norms are constructed, which will guide our interpretation of this perspective in this chapter. The first is the regulative pillar, formal and legally codified into explicit rules used to set clear parameters.

In South Africa, the advent and popularity of M&E with the establishment of the Department of Planning, Monitoring, and Evaluation was observed in 2010. Subsequently, the launch of the National Evaluation Policy Framework led to the formalisation of M&E practice and what is now formally known as the Government-Wide Monitoring and Evaluation system (Mouton 2010; Abrahams 2015). The formalisation of M&E in the government's national and provincial departments further led to a wide shift in non-governmental organisations, particularly those which provide public services, and private companies who are service providers to the government (Bornstein 2006; Abrahams 2015; Ngwabi, Mpyana and Mapatwana 2020). This is the approach that the regulative environment in the country uses to formalise results-based monitoring in its organisations.

The second pillar - the normative one - is defined by Scott (2008) as non-codified attitudes present in societies that can serve as normative expectations and attitudes that are gradually internalised by individuals, and eventually become accepted as the norms to which everybody is encouraged to conform. To this point, Powell and DiMaggio (1991) offer a practical explanation by stressing the importance of professional associations, foundations, and socialisation agents as drivers of change

in organisational forms. This underscores the role of Voluntary Organisations for Professional Evaluations (VOPEs), such as the South African Monitoring and Evaluation Association (SAMEA) and the African Evaluation Association (AfrEA), of shapers of 'results-based behaviour' among their individual and institutional members across the continent. Since their establishment, the two associations have influenced M&E practice (Goldman, Engela, Akhalwaya, Gasa, Mohamed and Phillips 2012; Basheka and Byamugisha 2015; Abrahams 2015; Cloete 2016; Kimaro and Fourie 2017). Following Powell and DiMaggio (1991) and Scott (2008), we would not be reaching too far to postulate that they also influenced the institutionalisation of results-based monitoring systems in many organisations.

Finally, Scott (2008) also asserts that while the foundational idea of the regulative pillar is 'conformity' to the rules and laws, the normative pillar recounts what is considered appropriate. Over time, the two are bound to influence individuals and groups at the cognitive level. Said differently, both rules and societal 'pressures' will eventually shape the way in which individuals think and solve problems. Arguably, this concept of cognitive influence is precisely the hope of organisational 'champions' of results-based monitoring systems. The anticipated level of institutionalisation of results-based monitoring will strengthen a culture of evidence-based decision-making.

However, contributing to our more profound understanding of organisational change from the perspective of institutional theory, Scott (2008) also provides a warning that depending on the nature and context of the pressure, responses to institutional pressures and expectations may range from passive conformity to active resistance. Perhaps Scott's idea of "the nature and context of pressures" may help us consider and understand our experiences of individual or group vested interests and fears that prevent them from embracing result-based monitoring systems.

The Study

Methodology

Purposive, convenient recruitment of M&E practitioners and programme managers who have been involved in building or strengthening an organisational monitoring system was used to identify participants for this study. The practitioners are conveniently picked from a list of M&E consultants and Data Innovators (Pty) Ltd clients. The practitioners ranged in experience from six to 21 years. The analysis is based on interview data from the 12 practitioners conducted between 09 January and 09 February 2022.

These practitioners represented experiences of results-based monitoring systems in a range of sectors, including Education and Skills Development, Health Sciences, Public Health, Governance, and Social Justice. The practitioners were interviewed

separately and consecutively to allow a cumulative understanding of the topic. This way, the researcher could adjust the interview guide or test emerging themes for subsequent interviews based on an analysis of previous ones. The interviews were unstructured, with a similar set of questions asked of each practitioner but allowing the scope for expansion, probing, or opening new issues. The questions investigated the participants' experiences of building and institutionalising results-based monitoring systems and their notions of `success factors' for institutionalising such a system within existing organisational norms and culture. Moreover, the study explored participants' views and descriptions of the complexities of organisational change during the process. For example, a common question asked to all participants during the interviews was, "What do you think are the factors that determine successful institutionalisation or adoption of a results-based monitoring system?"

All interviews were audio-recorded and fully transcribed for analysis. Narrative analysis, a qualitative method of investigation, focused on the research topic and analysis of the data collected from case studies and interviews. In this technique, the researcher identifies key findings and analyses them against the underlying topic and theories (Hemman and Vervaeck 2019). This method relies on the researcher's understanding of the background, setting, social, and cultural context of the research subjects, including the relevant theoretical frameworks. The analytical process focused on the meaning of the participants' experiences and phenomena, and included themes, contrasts, and theoretical explanations.

Findings

The study identified several interrelated themes detailing factors that determine successes and failures of the institutionalisation of results-based monitoring systems: (i) a decentralised (well-cascaded) knowledge, understanding, and appreciation of the value of a results-based monitoring system, (ii) adequate technical and user capacity among key staff, (iii) the explicit alignment and coordination of the results-based monitoring system within the broader organisational system, and (iv) the appropriate positioning of the results-based monitoring system as a learning process, and not a punitive measure. The four key factors are discussed in this section. The study also found that the explanatory power of theoretical perspectives of organisational systems varies with the context of the underlying organisation. The discussion sub-section outlines this relationship.

Decentralised Knowledge, Understanding, and Appreciation of Results-Based Monitoring Systems

Many practitioners in the study found knowledge of M&E and awareness of the objectives of results-based monitoring to be key factors of success for meaningful implementation. Respondents who observed significant success in the

implementation of a functional monitoring system attribute some success to staff awareness and working knowledge of M&E processes. These organisations were open to the sharing of information and were willing to accommodate M&E activities during daily operations, but there were also experiences of appropriate behaviour and support for meeting M&E requirements. Appreciation of the potential value of the results-based management to performance was said to have been an encouraging factor for management to embrace learning and the demand for performance data. It is essential to qualify this finding by clarifying that successful implementation was not observed among respondents who represented organisations where knowledge of RBM and M&E was concentrated among a few staff members. All respondents agreed that shared knowledge was more important than deepened knowledge concentrated amongst a 'handful' of team members. Two practitioners explained that when knowledge and appreciation are concentrated, the monitoring system tends to be owned and driven by a few individuals. It operates as a parallel system to the standard operating, accountability, and performance system. Even though the monitoring system can produce rich and valuable insights, these do not necessarily translate into learning or value for the organisation.

An example was provided by a respondent who represented the monitoring system of an organisation with a successful research and publication record. The respondent's monitoring system produces compelling monitoring insights that are disseminated widely amongst delivery partners. Although this monitoring system is highly productive, the uptake of the insights produced through programme monitoring has proven to be poor, resulting in little influence on decision-making. These practitioners attributed this observation to the fact that individuals who can make decisions for change in the organisation are not necessarily the same individuals who have an appreciation of the potential value of monitoring insights.

Respondents who experienced poor implementation of the results-based monitoring system also ascribed implementation failures to lack of awareness and understanding of M&E and the objectives of results-based management. As a result, in such organisations, staff who are meant to perform critical monitoring activities tend to show low participation when they do not understand the point of monitoring or its value to their work. Such staff members, therefore, default back to traditional routines.

To ensure shared understanding and appreciation of the value of the results-based monitoring system, almost all practitioners agree that a clear organisation-wide communication strategy before and during the change process is critical. One respondent emphasises the importance of treating the introduction and implementation of the results-based monitoring system as an organisational change project with a clear communication strategy for staff.

Technical and User Capacity

Another challenge regarding the institutionalisation of results-based monitoring systems is the low or entirely missing technical and user capacity required to perform monitoring and data processing activities. Practitioners define technical capacity as the skills or ability to perform basic monitoring activities such as data capturing, data cleansing, data translation, and data summaries. On the other hand, user capacity refers to the ability to interpret summarised information and monitoring reports and to facilitate reflections amongst decision-makers based on insights from monitoring reports. Although for some monitoring systems, the collation and processing of monitoring data can be entirely outsourced to an external service provider, it remains necessary for the users to possess the skills to interpret, reflect, and make decisions based on monitoring insights. Respondents admit that sophisticated monitoring systems include functional Information Communication Technology (ICT) or a digital platform used to automate most data collation and processing activities, thereby reducing turnaround time and decreasing the cost of monitoring. However, although such a system inherently reduces the technical capacity needed for manual processing, it requires investment in a technical capacity to use the digital platform.

Respondents attribute failure to sustain the functioning and use of the monitoring system to low technical capacity. To remedy or avoid this capacity 'pitfall,' practitioners suggest several contextually dependent solutions:

- Organisations with existing research capacity in their human resources can invest in basic M&E training to ensure in-house technical capacity.

- Organisations with no existing research or transferable analytical skills are advised to invest in the recruitment of M&E specialists to manage the technical and coordination functions of the monitoring process. In the latter scenario, the investment in recruitment can be augmented and offset by investment in a digital platform. In response to the scarcity of M&E capacity, contemporary monitoring practice has also encouraged outsourced and full-time secondment of practitioners from specialist M&E consulting companies to organisations with low technical capacity.

- For all contexts, respondents emphasised organisation-wide user training as utilisation as being a key component of the monitoring process. The practitioners who participated in the study stressed the requirement of critical decision-makers and facilitators thereof being sufficiently prepared for the use of monitoring outputs. Moreover, it is not sufficient to assume that the availability of monitoring insights will lead to uptake and use. In practice, this can be in the form of formal training. However, some respondents recommend the help of external experts to facilitate the implementation of the monitoring system in the early stages as a form of capacity transfer. This was said to be more effective than formal once-off training.

Alignment and Coordination of the Results-Based Monitoring System and the Organisational System

The extent to which proper arrangements for embedding new monitoring system processes within an existing operational and management protocol was identified as a critical factor for institutionalisation. In this case, respondents applied scrutiny to align proposed monitoring processes with existing operational and management procedures. Monitoring is intended as a management exercise and tool to ensure course correction during programme implementation. Practitioners argue that any activities performed as part of the monitoring process should not be divorced from the existing management protocol. They support the 'ideal' that monitoring should enhance and strengthen the managerial function. One respondent made the following remark: "There should never be parallel management systems. Monitoring is, by design, a management tool. Its objective is, and should be, to strengthen management by mainstreaming performance evidence."

Practitioners reported instances of decisions taken to change programme implementation approaches without consideration of monitoring insights, sometimes against the recommendations of monitoring reports. Another respondent reported changes in performance results (which have direct implications on what gets measured) without consultation with coordinators of monitoring personnel. Other examples include senior managers disregarding the role of monitoring in planning, or the misalignment between the programme monitoring framework and the human resource performance management framework. This results in the translation into parallel and unintegrated accountability systems. This misalignment leads to the 'strange' situation in which individual staff members can be rewarded for good performance while the organisation performs poorly against its targets.

Donor-funded organisations and jointly implemented initiatives are said to create an additional level of complexity to coordination and alignment. In these cases, the alignment challenge is not only between the results-based monitoring system and the existing organisational system, but also misalignment between the monitoring system of two organisations. In the case of a donor-funded organisation, the complexity is in aligning the monitoring and accountability system of the implementing organisation and the monitoring and accountability system of the funding organisation. In the case of jointly implemented initiatives, the challenge is in ensuring the alignment of the monitoring system of the two collaborating organisations.

Although there are varying levels of sophistication and complexities of misalignment in the different contexts, the challenges remain the same. This can be referred to as 'poor harmonisation' of systems. To remedy or mitigate this phenomenon, practitioners suggest that the design, introduction, and implementation of the results-based monitoring system needs to be intentional regarding the assessment

of disparities between systems and must set a clear plan for addressing identified conflict to maximise harmonisation. As such, implementing agents need to plan for potential misalignment and lack harmonisation between systems.

Positioning of the Result-Based Monitoring System as a Learning Process

A common misconception of the role of monitoring is the view and fear that monitoring is a punitive measure intended to 'police' performance of staff. Individuals perceive monitoring as a means to expose poor performance. As a result, participation in the monitoring process tends to be done begrudgingly and with fear. This 'atmosphere' of participation can limit meaningful learning. An example, as stated by several practitioners, is when staff are selective and preservative about the type, magnitude, and frequency with which they volunteer needed monitoring data, thereby stifling the quality of monitoring reports and insight. This results in the limiting of management's ability to draw meaningful conclusions from the organisation's performance. Practitioners describe this situation as the effect of mispositioning monitoring as a punitive auditing process with the primary objective of finding fault in performance. They emphasise that this perception can be overcome by offering appropriate training regarding the fundamentals of M&E to the staff. Practitioners who have experienced successful implementation of monitoring systems confirmed that one of the benefits of good M&E training for staff is the result of enlightenment and managed fear of scrutiny.

Some respondents believe that the fear of scrutiny and policing that can come with a strong monitoring system can be a fair conclusion and an appropriate assessment of the role of monitoring. They assert that in inherently bureaucratic organisations driven by high levels of competitiveness and poor culture of learning, a monitoring system can easily contribute to the existing culture of 'policing'. In these situations, practitioners recommend training intended as a form of advocacy targeted at the policy-making levels of the organisation. In this case, M&E and RBM training objectives should include the influence of performance management policy in favour of optimal learning.

Discussion and Conclusion

In this section, the study attempts to ground its findings in organisational change theory. It uses empirical data collected from practitioners to test and discuss the four perspectives of organisational change. It ascertains whether each of the four perspectives holds true and whether they can provide valuable insights to explain the challenges for the institutionalisation of results-based monitoring systems.

The Consequence of Change Theory Perspective

The consequence of change theory postulates that change failure is equivalent to 'organisational death' if the change includes changes to the core structures of the organisation. The study established that any introduction of the results-based monitoring system is tantamount to changes in articulating the organisation's goals, shifts in implementation, accountability mechanisms, and communications and marketing. The introduction of results-based monitoring is inevitably a shift in the organisational core structure.

When the consequence of change theory was tested against the experiences from respondents, the study found that this perspective holds selectively. Respondents from organisations that showed good levels of adoption of results-based monitoring reported instances of radical changes. The slowing down of the change process mitigated experiences of management discomfort and diminishing participation, as seen by performance reflections led by monitoring functions. Interestingly, the study also found that respondents and organisations who experienced radical changes were also the same organisations who (as part of the introduction of the monitoring system) mainstreamed evidence use and learning as core to their strategy. Examples of such organisations include ones that identify as 'Development Think-Tanks', 'Thought-leaders', and Innovation-driven organisations. There is significant reliance on systemically and accurately collected monitoring data for learning and innovation as the primary purpose in all these organisations. And thus, in this context, it should be expected that new evidence is likely to affect the organisations' core structures significantly.

Several of the remaining respondents reported a contradictory view. This group experienced the development and implementation of results-based monitoring systems as being intentional in the lack of change to the core structures of the organisations. Rather, the monitoring systems were designed to adapt to, match, and complement existing management systems and existing systems of decision-making and accountability. Therefore, organisational goals, operating patterns, authority mechanisms, and communication approaches were used as inputs to M&E framework design. In these organisations, the monitoring systems serve only the role of systematically collecting performance information, without altering operating norms, and availing insights for moments of reflection and decision-making as dictated by the management 'rhythm' of organisations. In this case, the monitoring function operates independently and parallel to other organisational procedures.

For such parallel systems, two distinct behavioural observations are made. The first observation is that management dynamically shifts mandate without due process, and often without communication to the monitoring function, and as such, the

monitoring function is always "playing catch-up". This leads to misalignment between the constantly changing mandate and what gets measured. Respondents describe this observation as a means for management to escape accountability. In other words, management overly shifts mandates to escape accountability. In the second observation, the parallel structure encourages management to use the monitoring function as a communication channel to report positive performance narratives. For this group of organisations, there is a strong expectation from management that the monitoring function should affirm the rhetoric of a positive performance story. This is particularly true for programme contexts where the introduction of the monitoring system is a function of donor or funder pressure and not necessarily a need for learning by management.

Structural Contingency Theory Perspective

Structural Contingency theory offers two arguments. The first argument is that demand for organisational change (demand for results-based management) is driven by changes in the market environment. The second argument is that organisations with more awareness and knowledge of the value of M&E invest more in the implementation of organisational change.

Most of the respondents found both arguments valid for implementing results-based monitoring systems. Respondents observed that results-based management and M&E are popular mechanisms to strengthen grant accountability and communicate progress among donors. As a result, it is generally difficult for implementing agencies to receive funding without demonstrating plans and capacity to conduct monitoring and evaluation. Even in instances where organisations do not have the technical capacity to manage a monitoring system, donors usually make the development of the M&E plan a condition for releasing funds, thereby compelling grantees to recruit an M&E partner. In these contexts, investment and resource allocation for the results-based monitoring system are easier to obtain. However, respondents argue that funder-pushed results-based monitoring systems can harm institutionalisation. They argue that, since the introduction of such a system is a function of pressure from funders, the implementation is primarily driven by the obligation to comply with funder requirements. However, the exception to this are instances in which the grantee organisation has an established good culture of evidence use. In this case, support and push from the funder can expedite and strengthen the institutionalisation of the results-based monitoring system.

Although it is generally agreed that good awareness and grounded knowledge of RBM and M&E makes the institutionalisation process 'smoother', knowledge alone is insufficient to deepen the institutionalisation. Respondents caution that "just because people know better, does not mean they will do better". They warn that although knowledge is necessary, structural dynamics like incentives, politics, and

values can influence behaviour more powerfully than knowledge and awareness can. They moderate the M&E capacity-building interventions that offer and focus solely on staff training and recommend that training interventions are more effective when other organisational factors are enabled.

Political Theory Perspective

It was confirmed across all cases that in each social programme or organisation, a coalition exists that is responsible for the critical 'contingency' of the whole. This grouping of the organisation is responsible for the core business of the programme or organisation and is referred to as the dominant coalition. The respondents agreed that any shift in the dominant coalition is tantamount to a shift in organisational strategy for or against the intended organisational change. One respondent explained that changes in the core workstreams of the education programme that they were implementing for two years always translated to changes in how monitoring was implemented across the programme. Dominant coalitions host the bulk of the monitoring effort and largely influence most of the monitoring processes. The coalition's resistance can easily lead to change or failure in the results-based monitoring system. Another response was that "chaos in the dominant coalition can lead to chaos in the entire change process".

The inverse of this political theory has also been proven to hold. Amongst respondents who experienced successes in institutionalising results-based monitoring systems, the influence and leadership of the dominant coalition in driving and normalising the use of monitoring insights proved to be effective in strengthening the monitoring system. Thus, when members of the dominant coalition are enlightened and technically capable, they lead the charge to build capacity across the rest of the organisation and champion M&E. Identifying and recruiting individual champions to act as ambassadors for the results-based monitoring system can be effective if the champions can influence the dominant coalition.

Institutional Theory Perspective

Finally, the institutional theory suggested that a results-based monitoring system can be more effectively institutionalised if its implementation is set as an official rule. i.e., endorsing and setting the implementation of the monitoring system as organisational policy deepens the change process. The study experienced inconclusive findings concerning this view, as the respondents were almost equally split between two categories of experiences. In one category of experiences, making reporting and reflection a compulsory activity with clear templates and guidelines assisted in the standardisation of the quality of reporting and ensured the conversion of insights to a decision. For example, an explicit agreement and contract between donors and grantees that obligates grantees to report progress has consistently improved

commitment to the implementation of results-based monitoring systems. However, it is important to note that this is only successful if other enabling factors, such as adequate capacity, understanding of the value of performance evidence, and clear and user-friendly guidelines, are in place. Thus, a rule on its own is not sufficient in the attainment of institutionalisation.

In the other category of experiences, it was found that an obligation to produce funder reports through standardised templates does not necessarily foster the adoption of the concomitant monitoring system. Instead, reporting is done purely as a compliance exercise that is misaligned to systems or organisational learning. Respondents emphasised that a rule to report consistently must be matched by consequence to the rule and appropriate incentive to adhere and participate meaningfully. A rule to monitor and report on performance is unlikely to be adhered to when the perception is that the consequence is absent or weak.

Furthermore, grantees need to have appropriate incentives beyond the written or legal obligation to monitor progress to avoid compliance-driven monitoring and reporting. An example of such an incentive can be the belief in the likelihood of improving and enhancing performance due to monitoring insights. Therefore, the rule must be accompanied by a clear sense of consequence and appropriate incentive to adhere and participate meaningfully.

Reference List

Barnett, W. P. & Carroll, G. R. 1995. Modelling internal organisational change. *Annual review of sociology*. 21(1), pp. 217–236. https://doi.org/10.1146/annurev. so.21.080195.001245

Bornstein, L. 2006. Systems of accountability, webs of deceit? *Monitoring and evaluation in South African NGOs. Development.* 49(2), pp. 52–61. https://doi.org/10.1057/ palgrave.development.1100261

Breakfast, N. B., Nomarwayi, T. & Dodd, N. 2021. Liberation Songs in Protest Actions: The Case of Service Delivery Protests in Nelson Mandela Metropolitan Municipality, Eastern Cape Province in South Africa. *African Journal of Development Studies* (formerly AFFRIKA Journal of Politics, Economics and Society). 2021(si2), pp. 143–160.

Burns, T. E. & Stalker, G. M. 1961. *The management of innovation.*

Clements, P. 2005. Monitoring and Cost-Effectiveness in Development Management. *Journal of MultiDisciplinary Evaluation.* 2(2), pp. 11–38. https://doi. org/10.56645/jmde.v2i2.118

DiMaggio, P. J. & Powell, W. W. 1991) *The iron cage revisited.*

Gosling, L. & Edwards, M. 2003. *Toolkits: A practical guide to planning, monitoring, evaluation and impact assessment.* Save the Children UK.

Hage, J. T. 1999. Organisational innovation and organisational change. *Annual review of sociology.* 25(1), pp. 597–622. https:// doi.org/10.1146/annurev.soc.25.1.597

Hage, J. & Powers, C. 1992. *Post-industrial lives: Roles and relationships in the 21ˢᵗ century.* Sage. https://doi. org/10.4135/9781483325828

Hannan, M. T. & Freeman, J. 1984. Structural inertia and organisational change. *American sociological review.* pp. 149–164. https://doi. org/10.2307/2095567

Herman, L. & Vervaeck, B. 2019. *Handbook of narrative analysis.* University of Nebraska Press. https://doi.org/10.2307/j. ctvr43mhw

Kusek, J. Z. 2010. *Making monitoring and evaluation systems work: A capacity development toolkit.* World Bank Publications.

Kusek, J. Z. & Rist, R. C. 2002. Building results-based monitoring and evaluation systems: Assessing developing countries readiness. *Zeitschrift* für Evaluation. 1, pp. 151–58.

Ladyman, J., Lambert, J. & Wiesner, K. 2013. What is a complex system? *European Journal for Philosophy of Science.* 3(1), pp. 33–67. https://doi.org/10.1007/ s13194-012-0056-8

Lahey, R. 2010. *The Canadian M&E System.* Washington, DC: World Bank.

Mackay, K. 2006. *Institutionalization of monitoring and evaluation systems to improve public sector management.*

Mapitsa, C. B. & Chirau, T. J. 2019. Institutionalising the evaluation function: A South African study of impartiality, use and cost. *Evaluation and Program Planning.* 75, pp. 38–42. https://doi. org/10.1016/j.evalprogplan.2019.04.005

Mashamaite, K. 2014. Public service delivery protests in a democratic South Africa: A dilemma for local municipalities. *Mediterranean Journal of Social Sciences.* 5(25), p. 231.

Mouton, C. 2010. *The history of programme evaluation in South Africa.* Doctoral dissertation.

Ngwabi, N. T., Mpyana, O. G. & Mapatwana, A. 2020. *Reflections from emerging evaluators in shaping Voluntary Organisations for Professional Evaluation capacity-building initiatives.* https://doi. org/10.4102/aej.v8i1.509

Pfeffer, J. & Salancik, G. R. 1974. Organisational decision making as a political process: The case of a university budget. *Administrative science quarterly.* pp. 135-151. https://doi.org/10.2307/2393885

Pfeffer, J. & Sutton, R. I. 2006. Evidence-based management. *Harvard business review.* 84(1), p. 62.

Salancik, G. R. & Pfeffer, J. 1978. A social information processing approach to job attitudes and task design. *Administrative science quarterly*, pp. 224-253. https://doi.org/10.2307/2392563

Scott, W. R. 2008. Lords of the dance: Professionals as institutional agents. *Organisation studies*, 29(2), pp. 219-238. https://doi.org/10.1177/0170840607088151

Sheth, K., Lorig, K., Stewart, A., Parodi, J. F. & Ritter, P. L. 2021. Effects of COVID-19 on informal caregivers and the development and validation of a scale in English and Spanish to measure the impact of COVID-19 on caregivers. *Journal of Applied Gerontology*, 40(3), pp. 235-243. https://doi.org/10.1177/0733464820971511

Simister, N. 2000. *Laying the Foundations: The role of data collection in the monitoring systems of development NGOs.* Occasional Paper, Centre for Development Studies, University of Bath, UK.

Swiss, J. E. 2005. A framework for assessing incentives in results-based management. *Public administration review*, 65(5), pp. 592-602. https://doi.org/10.1111/j.1540-6210.2005.00486.x

Vähämäki, J., Schmidt, M. & Molander, J. 2011. *Results based management in development cooperation.* Riksbankens Jubileumsfond: Stockholm.

Chapter 5

Acknowledgement

Chapter 6: A Systemic Lens to Indicator Development and Analysis

Jamie Robertsen

Introduction

Indicators are central to the development and implementation of monitoring systems. Depending on how the indicators themselves are developed and tracked, indicators can provide helpful information within a monitoring system. Equally, indicators can add a layer of cumbersome compliance to already overloaded monitoring and evaluation (M&E) teams.[22] Due to top-down approaches to selecting indicators, using indicators that are not fit-for-purpose, the expense associated with indicator tracking, low levels of use, and not accounting for complexity, indicators have become a controversial component of monitoring systems.

Box 1: What are indicators?

"Indicators are the quantitative and qualitative variables that provide a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of an organisation against the stated outcome" (Kusek and Rist 2004:65). Indicators:

1. Provide a simple and reliable way to measure achievement;

2. Reflect the changes connected to the intervention; and

3. Help assess performance against an objective.

In the field of M&E, indicators are often assessed on the extent to which they are Specific, Measurable, Attainable, Relevant, and Time-bound (SMART); Clear, Relevant, Economic, Adequate, Monitorable (CREAM); or Subjective, Participatory, Interpreted, Cross-checked, Empowering, or Disaggregated (SPICED). While there are other measures, these are the most common (World Bank, 2007). SMART and CREAM tell us what the indicators should look like, while SPICED gives us additional information on how the indicators should be used and developed (Lennie, Tacchi, Koirala, Willmore and Skuse 2011).

---

22   While this chapter focuses on monitoring specifically, where staff (practitioners), teams, or systems are referenced, the chapter notes these as M&E teams / M&E practitioners / M&E systems. This is because these functions are rarely separated in organisational structures or staffing.

Indicators provide managers and decision-makers with key information on the achievements, changes, and performance of their interventions (Kusek and Rist 2004). When this information is received timeously, decision-makers and managers can adapt and course-correct where needed. In the context of monitoring systems, indicators provide a framework for the monitoring system and guide the type and frequency of data collected.

Monitoring does not take place in a vacuum but a system. Interventions are complex and take place in different contexts characterised by multiple stakeholders with different interrelationships, levels of power, and contrasting views. Better understanding the systems in which monitoring occurs improves the practicality and depth of the approach to indicator development, analysis, and use.

With this in mind, this chapter uses a critical systems thinking lens to view indicator development, analysis, and use. Furthermore, this chapter aims to provide evidence and insight on indicators for African M&E practitioners. The chapter begins by outlining the challenges that practitioners face in the context of indicators, drawing on the experiences of African countries (Section 2). With a clear understanding of the challenges that M&E practitioners face, the chapter then grounds the discussion on indicator development and analysis, particularly the importance of meaningful measurement and use, in systems thinking (Section 3). Within a systems thinking perspective, the chapter focuses specifically on indicators in the context of complexity, boundaries, interrelationships, and contrasting perspectives (Section 3.4). The chapter concludes by looking at what a systems lens means for indicator use and influence (Section 4).

Challenges in the Development and Analysis of Indicators

Indicators are challenging and become a controversial piece of monitoring systems when: (i) they are developed in a way that does not take context and complexity into account, (ii) they are not useful, (iii) they are not developed in the context of the existing capacity or resources of a team or organisation, and (iv) they are seen as an external or top-down compliance measure. Focusing specifically on developing countries, a criticism of monitoring is that it emerged as a practice from countries with greater institutional capacity and more resources (Blaser-Mapitsa and Khumalo 2018).

Indicators can reduce reality to numbers, while reality is complex and nuanced (De Kool and van Buuren 2004). Furthermore, when indicators are not developed with the context in mind, they tend not to provide useful or needed information (UNAIDS). The following sentiment best sums this up: "It is important to keep indicators to the minimum and not to indulge in 'data greed', but once people are collecting the information, they should use it and act on it" (Silitonga et al. 2013:43).

In 2015, the International Labour Organisation (ILO) commissioned a meta-study to draw lessons from the evaluations and M&E systems that the organisation had developed and completed. Overall, the study found that there was inadequate measurement across the full results chain of the interventions, which minimised the extent to which the indicators could reflect a complete picture of the interventions (Lahey 2015). Similarly, in a study of the national M&E systems of Benin, South Africa, and Uganda, it was found that 45% - 52% of monitoring focuses only on activities and outputs, limiting the full view and potential impact of an intervention (Goldman, Olaleye, Ntakumba, Makgaba, and Waller 2021).

A challenge that is frequently brought up in the literature, but does not relate to indicators specifically, is low levels of monitoring capacity (Dipela and Mohapi 2021), which is particularly true for smaller organisations and under-staffed M&E units. Holvoet, Gildemyn, and Inberg (2012) assessed the M&E systems of 20 aid-dependent countries in sub-Saharan Africa. Coordination and oversight were categorised as satisfactory in 25% of the countries studied, and partially satisfactory in 75%. Of the same twenty countries, 15% had excellent capacity building plans, 5% had satisfactory plans, 60% had partially satisfactory plans, and 15% had weak plans. These findings point to considerable capacity and management constraints. Goldman et al. (2021) mirror these findings by highlighting that in Benin's M&E system, a high turnover of staff, low levels of institutional memory, and limited resources contribute to capacity constraints in the system; while in Uganda, the government's M&E system is understaffed and the M&E capacity levels are low.

The 2015 study commissioned by the ILO also noted that indicators often served an administrative purpose, were used to report on activities, and were used to reach milestones that would result in a release of funds (Lahey 2015). This aligns with a commonly held view that indicators are tools of compliance and accountability (Dipela and Mohapi 2021) and are a means for funders to exert power (De Kool & van Buuren, 2004). When indicators are used as an external donor or internal top-down compliance tool, low stakeholder engagement and participation levels often result in indicators not wholly owned by an organisation (Kusek and Rist 2004). Low levels of buy-in negatively impact the extent to which indicators are meaningful and can be measured and used. In Benin and Uganda, for example, donors play a prominent role in funding government-level monitoring. In both countries, monitoring is seen as compliance-driven, staff have reporting fatigue, and managers only use 50% of the monitoring data collected (Goldman, Olaleye, Ntakumba, Makgaba, and Waller 2021). In the study mentioned above on the M&E systems of 20 countries in sub-Saharan Africa, 45% showed partially satisfactory use of monitoring data, and 55% showed weak use of monitoring data (Holvoet, Gildemyn and Inberg 2012).

While these challenges are numerous, there are several strategies to mitigate the extent to which they affect the development of meaningful and useful indicators. An

important starting point to mitigate these challenges is the adoption of a systems thinking lens to indicator development, analysis, and use. A systems thinking approach grounds indicator development in what a system is in reality rather than an idealised version. Once a system's complexities, boundaries, and interrelationships are understood, practical approaches to indicator development can be adopted.

Systems Thinking as a Lens to Improve Meaningful Measurability

The systems in which M&E practitioners operate are inherently complex in that interventions, and the systems in which they operate, have attributes that are difficult to understand and conceptualise; there are often multiple perspectives and interests. The intervention within the system is non-linear, and there are high levels of uncertainty.

THE WAY WE INTERACT WITH A SYSTEM IS INFORMED BY:

Values and motivations

Knowledge base

Power structures

IN A SYSTEM, WE NEED TO:

Understand interrelationships

Engage with contrasting perspectives

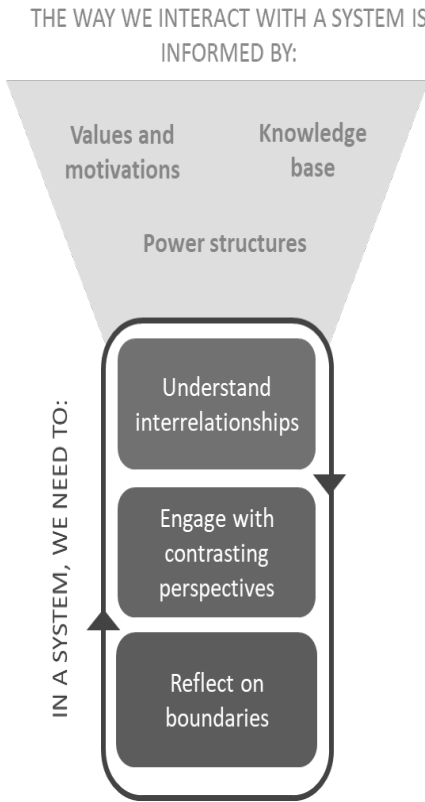Reflect on boundaries

Figure 6.1:   Critical systems thinking

Source: (Reynolds, 2014)

There are three critical elements to understanding and conceptualising a system, as shown Figure 1: (i) reflecting on the boundaries in which you are working, (ii) understanding the interrelationships within a system, and (iii) engaging with contrasting perspectives. Furthermore, values and motivations, a knowledge base, and power structures inform how stakeholders interact with a system (Reynolds, 2014).

The work done by Reynolds (2014) on the use of critical systems thinking in evaluations is adapted and used as the framework through which this chapter applies a systems lens to indicator development and analysis. Table 1 shows the stakeholders, stakes, and stakeholding issues that should be considered when assessing a monitoring system for who gets what, who owns what, and who does what.

This framework provides a way to understand the boundaries and stakeholders (and their interests) within a monitoring system.

Table 6.1: Reynolds, *Boundary judgements for informing a monitoring system*, 2014

Sources of influence in a monitoring system	Boundary judgements informing a system of interest for an intervention (policy, programme, project)			
	Stakeholders	Stakes (specific interests)	Stakeholding issues (key problems)	
Who gets what? *Sources of motivation*	1. Primary users *Who will receive the data from the indicators collected and analysed?*	2. Purpose *What do the primary users need the data for? How will they use it?*	3. Measure of success *What do good, meaningful, practical indicators look like in this context?*	
Who owns what? *Sources of control*	4. Decision-makers *Who has the resources and influence to develop, collect and analyse the indicators for the monitoring system?*	5. Resources *What resources are available for indicator development, data collection, and analysis? What can practically be done with these resources?*	6. Decision environment (accountability) *Who is ultimately responsible (and accountable) for using indicator data?*	*The involved*
Who does what? *Sources of knowledge*	7. Experts *Who will be responsible for developing indicators, collecting data and analysing the collected data? Who will manage this process?*	8. Expertise *How much capacity is available in the organisation or unit outside the direct management and team? What can practically be done with this capacity?*	9. Guarantor *Who will provide additional advocacy and leadership support to the indicator development, analysis, and use process? Will the work feed directly into decision-making structures?*	

In answering these questions, the M&E practitioner will know:

- The users of their data and the needs of these users to develop a practical, meaningful set of indicators.

- The key decision-makers with the resources to develop and analyse indicators, and what resources are available for this undertaking. This allows the M&E practitioner to develop a monitoring system and indicators within the resource parameters of their team or organisation.

- Who is responsible for developing and analysing the indicators, and how much additional capacity and support is available for this endeavour? This allows the M&E practitioner to develop a monitoring system and indicators within the capacity constraints of their team or organisation.

Once the bounds of the system are understood, it is possible to make indicator decisions that are meaningful and fit-for-purpose. The first of these decisions is whether the approach to developing the indicators will be results-based or informed by a logic model, or whether the approach will include standardised indicators (see Box 2).

Box 2: The pros and cons of using standardised indicators

Various standardised indicators are available – ranging across multiple sectors and multiple levels of government – from national to regional and multilateral. These include, for example, (i) sets of indicators aligned with country-level national development and poverty alleviation plans, (ii) indicators linked to the Sustainable Development Goals (SDG), (iii) bilateral donor indicators such as USAID's standard indicators, and (iv) the indicators in the World Bank's Rural Development Handbook (World Bank 2007).

The pros of using standardised indicators are that the effects of projects, programmes, and policies can be aggregated and compared; the cost of building unique measurement systems is reduced; and there can be harmonisation across donor requirements (Kusek and Rist 2004). Primarily, the pros tend to favour donors and governments in that standardised indicators enable them to aggregate the impacts of their interventions.

The cons of standardised indicators typically relate to the interventions (projects, programmes, policies) themselves. The use of standardised indicators amplifies the feeling of a top-down or imposed approach to indicators and monitoring; there is less participation and ownership; and if an organisation has multiple funders, it can lead to too many, sometimes competing, indicators (Kusek and Rist 2004; Moreno Pires, Magee and Holden 2017).

Indicators Within a Results-Based Management System that is Grounded in the Practicalities of Stakeholders, Resources, and Capacity

The most promoted approach to monitoring is results-based management because it focuses on monitoring as part of a broader decision-making toolkit. It focuses on both the process and implementation aspects of an intervention and its results (Lamhauge, Lanzi and Agrawala 2013). A results-based management approach to

monitoring and indicator development typically begins with articulation of the logic of an intervention, either in the form of a logical framework, results chain, or theory of change.

Articulation of the logic of the intervention and its intended outcomes allows common understanding of the processes to be undertaken and affords a view of what success would look like. This is shown in Table 2, which summarises the steps for establishing an M&E system (Kusek and Rist 2004), and the steps to monitoring a programme, as set out by the Donor Committee on Enterprise Development (DCED) (Sen, Kessler and Loveridge 2018). Once this has been articulated, indicators are developed to align with the logic of the intervention, user and decision-maker needs, and resources and capacity available in the system.

Table 6.2: Indicators in the context of monitoring systems and monitoring programmes

Ten steps to a results–based M&E system		Steps to monitoring a programme	
1	Conduct a readiness assessment	1	Articulate the results chain (theory of change)
2	Agree on outcomes to monitor and evaluate	2	Define indicators of change, other information needs
3	Select key performance indicators to monitor outcomes	3	Measure attributable change
4	Set baselines and gather data on indicators	4	Capture wider changes in the system
5	Plan for improvement – selecting results targets	5	Track costs and impact
6	Monitoring for results	6	Report costs and results
7	The "E" in M&E – Use evaluation information to support a results–based management system	7	Manage the system for results measurement
8	Report the findings	*(Sen, Kessler and Loveridge 2018)*	
9	Use the findings		
10	Sustain the M&E system within the organisation		

(Kusek and Rist 2004)

Kusek and Rist (2004:65) argue that "indicators should be developed for all levels of the results-based M&E system, meaning that indicators are needed to monitor progress with respect to inputs, activities, outputs, outcomes, and goals. Progress needs to be monitored at all levels of the system to provide feedback on areas of success and areas in which improvement may be required". Within this approach, input indicators measure the resources needed for the intervention, process indicators measure the implementation of the activities of the intervention, output

indicators measure the effectiveness of implementation, and outcome indicators measure the extent to which an intervention met its objectives (Huovila, Bosch and Airaksinen 2019).

Using a results-based management approach addresses several indicator-related challenges. It provides implementers and decision-makers with a fuller picture of the intervention, and the developed indicators are aligned to the objectives of the intervention. However, it is important to note that the indicators developed from a logical framework or theory of change should still be bound to the practical feasibility of their collection, analysis, and use. Prioritisation is central to developing indicators that will be tracked. In a study conducted by Holvoet and Renard (2007), it was found that nine of the 20 countries' M&E systems in sub-Saharan Africa had long lists of indicators, with Ethiopia being the highest at more than five hundred. In a later study, Holvoet, Gildemyn, and Inberg (2012) found that of the same twenty countries, 40% were considered weak in indicator priority setting, 40% were considered partially satisfactory, and 20% were considered satisfactory.

Box 3 highlights the importance of indicator prioritisation.

Box 3: Lessons learned from an assessment of the M&E systems of social interventions in Ghana, Kenya, Moldova, and Mozambique

Background

A 2014 study commissioned by Oxford Policy Management (OPM) looked at monitoring systems and developed key lessons on Social Safety Net (SSN) interventions in Ghana, Kenya, Moldova, and Mozambique. The interventions were: i) the Livelihood Empowerment Against Poverty (LEAP) cash transfer programme in Ghana; ii) the National Safety Net Programme (NSNP) in Kenya, which brings together five existing cash transfer programmes; iii) the Ajutor Social programme in Moldova, which provides cash benefits to applicants who fall below a certain income threshold; and iv) four different SSNs in Mozambique, including the country's largest cash benefit intervention, Programa de Subsidio Basico (PSSB).

The study used a supply-and-demand model as a framework for analysis. Based on this, the key conclusion on M&E systems is that when effective, they address both supply-side and demand-side factors. On the supply side, these systems provide reliable information, and on the demand side, these systems create the demand for and use data to support the SSN interventions.

Findings

Two key lessons were apparent, looking specifically at indicators in the SSN interventions in these four countries. The first was that indicators should be defined based on the information needs of the SSN intervention. The second was that arriving at both useful and feasible indicators required a process of prioritising, refining, and organising as an iterative process.

Indicators should be defined based on the information needs of the intervention

Based on a review of SSN interventions in the countries mentioned above, it is clear that indicators are useful and are used when they are demand driven. Information needs were found to depend mainly on i) the intervention's objectives, its theory of change, or results framework, and ii) the needs of different actors and stakeholders.

Indicators are developed through an iterative process of prioritising, refining, and organising

All four interventions undertook a process of extensive mapping of information needs and indicators. In Ghana, this initial consultation process resulted in more than one hundred indicators, and in Mozambique, the number was closer to two thousand. This initial process of consultation was central to building buy-in.

Once this long list of indicators was developed, more consultations and workshops were held with the system's main users to prioritise indicators. Key considerations in this prioritisation process were the accessibility of data for the indicators and identifying data sources.

In refining the indicators, the key criterion was the extent to which the proposed indicators were Clear, Relevant, Economic, Adequate, and Monitorable (CREAM).

Finally, the prioritised and refined indicators were organised based on their use. For example, in Ghana and Moldova, the indicators were organised into indicators for programme operations and management and indicators for results (Attah et al. 2014).

One of the first steps towards developing meaningful indicators is to ensure that the indicators are useful and fit-for-purpose. Meaningful indicators reflect the objectives of the intervention, are precise enough to reflect the change, and are practical and cost-effective (Kusek and Rist 2004). The development of meaningful indicators provides a complete view of an intervention and ensures that the indicators developed are practically feasible to implement and can be measured.

Indicators that Take Stakeholder Consultation, Management, and Contrasting Perspectives into Account

Participatory indicator development and analysis mitigates several of the challenges related to indicators and is firmly grounded in a systems lens that advocates for understanding stakeholders, their interests and influence, and for the consideration of contrasting perspectives for a more realistic view of the system. Increased participation can elicit unique insights that provide contextual information about interventions and bring the voices of beneficiaries to the fore.

Increased participation means that indicators provide a fuller picture of an intervention, that the development of indicators is bounded in practical feasibility, and that indicators are not seen to be imposed. These all contribute to a sense of ownership for the implementers of the intervention, its stakeholders, and its beneficiaries. In addition to a greater sense of ownership, a participatory approach can improve the quality of indicators produced and the quality of data collected (Kusek and Rist 2004). An example of the benefits of participatory indicator development is provided in Box 4.

Box 4: Participatory indicator development in the Kalahari, Botswana

Background

To develop meaningful environmental sustainability indicators that organisations, funders, decision-makers, and communities (in this case, pastoralists) could use, a team of researchers sought to develop a set of indicators for land degradation in the Kalahari in Botswana. Land management is central to environmental sustainability, and this work was undertaken because conservation work relies heavily on standardised indicators that are not necessarily useful to pastoralists.

The researchers undertook a four-step process: i) develop methods that can effectively integrate local and scientific knowledge on environmental sustainability indicators; ii) identify environmental sustainability indicators from pastoralists and available literature on the region; iii) evaluate this indicator knowledge qualitatively in community-level focus groups, and quantitatively using ecological and soil-based sampling; and iv) test the assumption that it is not possible to have both meaningful participation and scientific rigour. The researchers selected three sites where land degradation is a problem and where there were differences in biophysical and cultural settings.

Indicator Development

Through interviews, pastoralists were asked to identify signs on the land that would indicate land degradation over the long term. The pastoralists were asked to discuss which signs would appear first and might serve as an early warning. They were also asked to assess potential indicators based on how accurate they would be and how easy they would be to monitor and use. The researchers also used oral histories to understand historical environmental changes better.

Once a list of indicators was developed, these were monitored through ecological sampling and discussed in focus groups in the communities. A sample of these indicators is shown in the table below.

Results

Not all the indicators developed by the pastoralists were statistically significant. However, the researchers found the pastoralists to be a rich source of environmental sustainability and land degradation data. What the researchers found particularly useful was that the indicators were very localised and therefore more useful to the pastoralists than more standard environmental indicators. The researchers were also able to extend their indicator base beyond soil and vegetation. The pastoralists highlighted other important indicators of land degradation not traditionally used: (i) livestock spending more time eating bushes and foraging further from water points, and (ii) increased expenditure on goods typically sourced from the land.

Overall, local knowledge was more holistic in its development of indicators. The pastoralists were involved in developing these indicators and chose them based on accuracy and ease of use. The pastoralists, therefore, have some sense of ownership of the indicators and can continue monitoring their land.

Indicators considered accurate to and easy to use by pastoralists, showing evidence from literature and empirical testing

	Indicator	Supported by literature?	Site 1	Site 2	Site 3	
	Decreased grass cover	Yes	Significant	Significant	Significant	
	Increased abundance of grass unpalatable to cattle	Yes	Significant	Significant	Significant	
	Decreased abundance of grass palatable to cattle	Yes	Significant	Significant	Significant	
	Decreased availability of thatching grass	No literature	N/A for this site	Not significant	Significant	
	Decreased abundance in medicinal plans	No literature	Insufficient data	Insufficient data	Insufficient data	
	Decreased abundance of trees	No	Not significant	Not significant	Not significant	
	Stunting of trees and bushes	No	Not significant	Not significant	Not significant	
	Tree canopy die-off	No literature	Significant	N/A for this site	N/A for this site	
	Increased abundance of *boscia albitrunca*	No literature	Significant	N/A for this site	N/A for this site	
	Decreased abundance of *grewia flava*	Yes	Not significant	Insufficient data	Insufficient data	
	Increased abundance of *acacia mellifera*	Yes	Significant	N/A for this site	Significant	
	Decreased vegetation cover / increased bare ground	Yes	Significant	Significant	Significant	
	Decreased soil organic matter content	Yes	N/A for this site	Significant	Significant	
	Increased soil looseness	No	Not significant	Not significant	Significant	
	Increased density of cattle tracks	No literature	Not significant	Not significant	N/A for this site	

(Reed, Dougill and Baker 2008).

In addition to the value of participatory development of indicators, indicators can also be assessed from the perspectives of a range of different groups, which can provide a more nuanced view of the intervention, as shown in Box 5.

Box 5: The importance of disaggregating indicators

Background

The United States Agency for International Development (USAID) Resilient Waters Programme is a five-year programme that focuses on building communities' socioecological resilience in the Limpopo and Okavango River basins. The Programme aims to achieve this objective by working with transboundary river basin organisations and transfrontier conservation areas, improving access to water, sanitation, and hygiene, improving livelihoods, and supporting biodiversity and conservation.

The Programme's Monitoring, Evaluation, and Learning (MEL) team developed a composite indicator to measure resilience in the region. In doing so, Resilient Waters focused on the four capacities of resilience:

- Absorptive capacity: The ability of households to respond to shocks. Central to this is whom communities can turn to when there has been a shock or stressor – individuals, communities, and institutions.

- Anticipatory and adaptive capacity: The ability of a household to plan to prevent the negative results of future shocks.

- Transformation capacity: This capacity is concerned with existing power relations and building new social dynamics that shift resilience and reduce vulnerability to shocks.

The Importance of Disaggregating Data for Planning

In collecting data for this composite indicator, and to better understand absorptive capacity (who people turn to in times of crisis), the MEL team investigated the role of institutions in supporting communities. The baseline survey found that 56% of respondents reported that they know at least one organisation that they can turn to for help if their household is adversely affected by a shock or stressor that they cannot manage themselves.

When disaggregating the data, a more nuanced picture emerged:

- While 60% of males reported having at least one organisation to turn to, 53% of females reported having at least one. This highlights different levels of access to and trust in organisations across genders. From a planning perspective, Resilient Waters needed to develop a consistent and far-reaching approach to gender equity and social inclusion.

- Turning to youth (younger than 35) and non-youth (older than 35), there was no significant difference in whether people have an organisation to turn to for help. There was, however, a difference in outlook on whether these organisations could help. Youth tended to be more optimistic that these organisations could help (an average rating of 3.4 out of 5) compared to non-youth (an average rating of 2.8 out of 5). From a planning and programming perspective, this finding highlighted that youth are an effective, more trustful entry point for organisations working in these basins.

(USAID Resilient Waters 2019).

Indicators that Shed Light on Multiple Levels and Scales of an Intervention

Meaningful indicators should, (i) allow their users to understand implementation better (process indicators), (ii) serve as an early warning system for any potential challenges, and (iii) improve their users understanding of the results of an intervention (outcome indicators) (Kusek and Rist 2004). Process indicators are concerned with inputs, activities, and outputs. Outcome indicators are concerned with outcomes and impacts. Including both sets of indicators provides a fuller picture of an intervention.

Process indicators are more easily measured and, therefore, more commonly used and requested by funders. These indicators are helpful because they reflect what happened, but they do not provide information regarding the extent to which what has happened has led to any tangible results (Lamhauge, Lanzi and Agrawala 2013). For example, a process indicator would tell us how many schools have been built (output level), but not how many children are attending the school, or in fact, receiving quality education from the schools (outcomes levels) (Lamhauge, Lanzi and Agrawala 2013). There are key benefits to developing meaningful indicators at both levels, as highlighted in the example provided in Box 6 below.

Box 6: Developing outcome indicators to understand better the effects of sports-based HIV/AIDS interventions in South Africa

Background

Sports programmes are a common intervention for building awareness and responding to HIV/AIDS. These programmes tend to be aimed at young people and specifically focus on developing confidence and supporting health and healthy behaviours in the context of HIV/AIDS. Sports-related HIV/AIDS interventions in South Africa have traditionally focused on output level indicators as a source of monitoring which tends to limit the understanding of the effectiveness of these programmes. With this premise in mind, Maleka (2017) sought to develop standardised outcome indicators that could be used to understand the results of sports-related HIV/AIDS programmes better.

Results

Based on consultations with key stakeholders and organisations, a list of 51 outcome indicators was developed, categorised by the type of results sought. In this context, one of the purposes of developing generic outcome indicators was to lower the time burden on local organisations to develop these indicators. The indicators reflect the realities of these interventions, and local organisations are therefore able to select the indicators that are most applicable to them.

One of the categories of outcome indicators developed focused on stigma and discrimination. The table below outlines a set of generic outcome indicators that would show reduced stigma and discrimination towards people living with HIV/AIDS (PLWHA) (Maleka 2017:16). Output or process indicators for stigma and discrimination would typically focus on the activities associated with the intervention. For example, sports-related programmes could use education to reduce stigma and discrimination towards PLWHA. The process indicators would then focus on the extent to which the education took place and what participants learned from the activities. Outcome indicators would then build on this by providing insight into changes in participants' behaviours when it comes to stigma and discrimination.

Generic anticipated outcomes	Generic outcome indicators
Reduced HIV stigma and discriminatory attitude towards PLWHA.	Percentage of participants who report an accepting attitude towards PLWHA. Percentage of participants who are willing to talk, care for, and identify with someone who has HIV/AIDS.
Positive intention to communicate about HIV/ AIDS with peers and family.	Percentage of participants reported intention to communicate with someone outside of a programme about HIV/AIDS.
Increased percentage of HIV positive participants who report that they are comfortable disclosing their HIV status.	Percentage of HIV positive participants who express a positive ability to feel comfortable to disclose their HIV status to their sexual partners or any person they trust.

If used, the outcome indicators developed will provide a fuller picture of the effectiveness of sports-related HIV/AIDS programmes in South Africa. However, the author does highlight that reporting on these indicators would require more time and planning. Therefore, the indicators developed should be adapted to fit within the resource constraints of the implementing organisations.

(Maleka 2017).

Like the argument for using both process and outcome indicators, quantitative and qualitative indicators can provide a fuller picture of an intervention. For example, if the intervention is an employment programme, quantitative indicators can measure how many jobs were created. However, this does not indicate the quality of the jobs or what was or was not effective in the intervention. Qualitative indicators give a monitoring system explanatory power. They can shed light on perceptions of the programme among beneficiaries, unexpected changes, and the sustainability of the programme's changes, as shown in Box 7 (Sen, Kessler and Loveridge 2018).

Box 7: The use of qualitative and quantitative indicators in climate adaptation programmes

Background

Lamhauge, Lanzi, and Agrawala (2013) conducted a study on the use of indicators to measure climate change adaptation. The study assessed M&E frameworks used by development cooperation agencies and looked at the total frameworks of 106 projects.

Findings

Central to developing M&E systems in the context of adaptation is understanding that the context is complex, and that attribution is difficult because there are lags between interventions and outcomes. The authors further note that while quantitative indicators can and should be used, there is also a need for a more subjective, qualitative understanding of an intervention. For example, within the context of adaptation policy and administrative management, a policy being introduced (quantitative indicator) does not mean that it is effectively applied or mainstreamed (qualitative indicators).

Training on climate change adaptation is a popular intervention in adaptation programmes. Quantitative indicators can provide information on how many people were trained or if the training occurred, but qualitative indicators are required to understand the longer-term changes. Examples of qualitative indicators that can be used in conjunction with quantitative measures include the percentage of trained policymakers who applied the information learned, and the proportion of people who feel that they are prepared for natural shocks and stressors due to the training received.

(Lamhauge, Lanzi and Agrawala 2013).

What Does a Systems Lens Mean for Influence and Use?

Using indicators means that they are consistently tracked and reflected on and used to make decisions that relate to the implementation and programme direction. In the absence of use, indicators are rendered meaningless. There has been considerable research on use in the context of evaluation where four key types of use have been identified:

- **Instrumental use**: The direct use of evaluation results to improve programmes and provide information for decision-making.

- **Conceptual use**: The use of evaluation results to bring about improved understanding and new ways of thinking.

- **Symbolic use**: The use of evaluation results to legitimise decisions that have already been taken.

- **Process use**: The indirect use of the process of evaluation to change procedures, behaviours, and organisational culture (Weiss, Murphy-Graham and Birkeland 2005).

In the context of evaluation, use can be promoted during, (i) the implementation of an evaluation by improving evaluation quality, credibility, relevance, and timeliness; and (ii) in the decision-making setting by ensuring that the evaluation addresses information needs, that there is a favourable political climate in place, and that an organisation is receptive to the evaluation (Weiss, Murphy-Graham and Birkeland 2005). Expanding this view to monitoring and indicators would essentially mean that if indicators are SPICED, they are more likely to be used. While this is true to an extent, this view does not sufficiently take the context of use into account. In their book, *Monitoring and Evaluating Social Programmes in Developing Countries*, Valadez and Bamberger (1994) provide a broader view of use, which considers the decision-making context. The authors note that the following steps are essential in improving monitoring use:

- Making sure that data collection is well-timed and aligned to decision-making;

- Improving the quality of the data;

- Involving stakeholders;

- Building relationships to enhance people's incentives to participate and help;

- Identifying intended users;

- Improving communication and dissemination;

- Institutionalising responses to reporting;

- Building a learning culture within an organisation;

- Defining stakeholders and their information needs;

- Developing a strategy for working with different stakeholders;

- Integrating indicator development and monitoring into the implementation cycle; and

- Communicating clearly and consistently.

Hatry (2008) further highlights that indicator data will be used more if, (i) outcome data is provided in a timely and frequent manner, (ii) data is disaggregated by demographic and service characteristics, (iii) current data is compared to previous data to give it more meaning, (iv) explanations of unexpected findings are provided, (v) the data is effectively summarised, and (vi) reporting is user-friendly.

Many of the steps mentioned here have been addressed in prior sections, particularly in discussing the importance of participatory indicator development, the development of meaningful indicators, and being cognisant of the practical limitations of monitoring indicators.

134

Taking a systems approach to indicator use would extend the aforementioned conceptions of use beyond individual and organisational characteristics to a better understanding of the underlying processes and interrelationships that influence the use of any data (Mark and Henry 2004). Influencing indicator use or organisational and system change at any level is difficult because change does not occur in a vacuum and depends on the individual, interpersonal, and collective changes (Appleton-Dyer, Clinton, Carswell and McNeill 2012).

Going back to Reynolds (2014) and Figure 1 of this chapter, taking a system view on what happens to indicators once they are developed, collected, and analysed goes beyond conceptions of use to conceptions of influence. Influence in a system is a factor of values and motivations, power structures, and knowledge bases. Sources of influence in a system are, therefore, (i) motivation (who gets what?), (ii) control (who owns what?), (iii) knowledge (who does what?), and (iv) legitimacy (who is affected by what some people get, and who is potentially marginalised?). Therefore, it is critical to understand sources of influence and the interrelationships within the system of an intervention to influence change of any kind. In this context, they would influence the use of indicators and the information provided by them.

Reynolds' (2014) work on critical systems thinking builds on Mark and Henry (2004), where a framework of influence to consider change processes was developed. The framework was initially developed to influence evaluation use but has been adapted here to influence indicator data use, as shown in Figure 3. This adaptation is possible due to both monitoring indicators and evaluation providing data, and the goal of both is for this data to be used for the improvement of an intervention.

In using this framework, and drawing on Section 3, M&E practitioners need to understand the context in which they are working, the attributes and motivations of influential individuals or groups in their system, the mechanisms of change in their system, and to provide information that is responsive, credible, well-communicated, and timely. With this foundation, the framework highlights that indicator use can be influenced through cognitive mechanisms that provide consistent information, highlight the benefits of indicator use for improved decision-making, appeal to organisational norms, and are clear on how the indicators can be used. Additionally, the framework allows for motivational mechanisms which provide incentives and appeal to personal goals, as well as behavioural mechanisms which build on cognitive and motivational mechanisms to start, stop, or continue with a policy, project, or programme based on the available evidence.
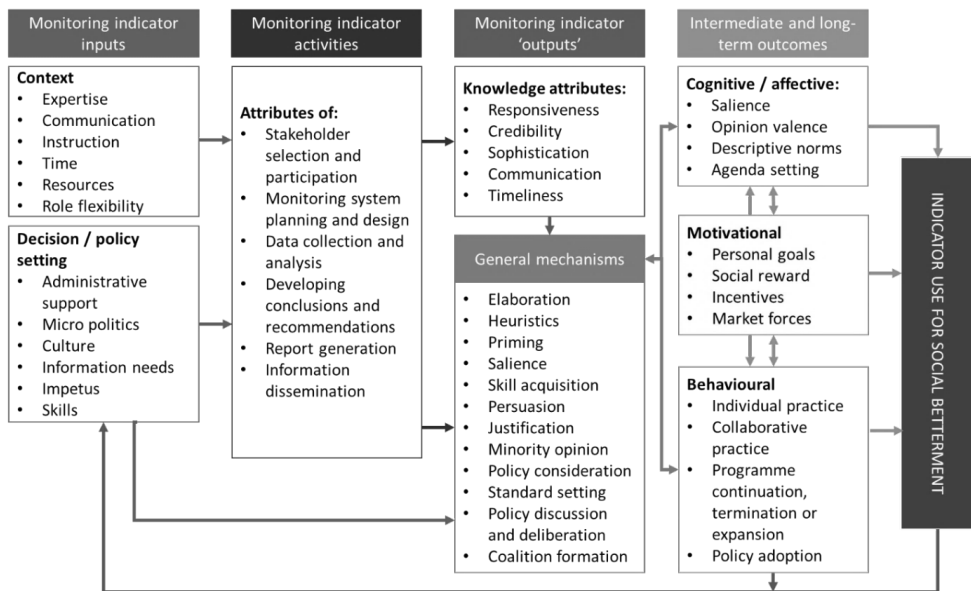
Contingencies in the environment: Competing processes, facilitating factors, and inhibiting conditions

Figure 6.3: Mark and Henry, Schematic theory of indicator data influence, 2004:46

Conclusion

Indicators are central to the development and implementation of monitoring systems and can provide useful information for interventions. When this information is received promptly, it allows decision-makers and managers to adapt and course-correct where needed. However, indicators and the monitoring systems within which they exist do not take place in a vacuum. To address the contextual challenges that indicators can pose, it is vital to take a systems view of indicator development, analysis, and use that is grounded in reality, rather than what should be.

In adopting a systems approach to developing indicators, M&E practitioners can understand the users of their data and their need to develop a practical and meaningful set of indicators. The need for key decision-makers who have the resources to develop and analyse indicators, as well as define the resources available for this undertaking, are made clear. This allows the M&E practitioners to (i) develop a monitoring system and indicators within the resource parameters of their team or organisation; (ii) determine who is responsible for developing and analysing the indicators; and (iii) identify how much additional capacity and support is available for this endeavour. This knowledge allows practitioners to develop indicators that are meaningful and fit-for-purpose.

This conceptualisation of indicators during indicator development provides a strong foundation for indicator use. The requirements and constraints of the system are understood, and the indicators that are developed should, therefore, be intrinsically designed for use. A systems lens takes this one step further, arguing that developing good indicators is vital for use but is not the only determinant because systems are characterised by underlying processes and interrelationships that influence indicator use. Therefore, indicator use can be influenced through cognitive, motivational, and behavioural mechanisms of change at the individual, interpersonal, and collective levels.

Reference List

Appleton-Dyer, S., Clinton, J., Carswell, P. & McNeill, R. 2012. Understanding evaluation influence within public sector partnerships: A conceptual model. *American Journal of Evaluation*, 33(4), pp. 532-546. doi:10.1177/1098214012447672

Attah, R., Barca, V., MacAuslan, I., Pellerano, L., Ragnos, L., Riemenschneider, N. & Simon, C. 2014. *How to move beyond the impact evaluation trap? Challenges and solutions for setting up of comprehensive M&E systems for social safety nets.* Oxford: Oxford Policy Management. https://www.academia.edu/31894737/How_to_move_beyond_the_impact_evaluation_trap_Challenges_and_solutions_for_the_setting_up_of_comprehensive_M_and_E_systems_for_Social_Safety_Nets.

Blaser-Mapitsa, C. & Khumalo, L. 2018. Diagnosing monitoring and evaluation capacity in Africa. *African Evaluation Journal*, 6(1). https://doi.org/10.4102/aej.v6i1.255

De Kool, D. & van Buuren, A. 2004. Monitoring: Functional or Fashionable? *Society and Economy*, 26(2/3), pp. 173-193. http://www.jstor.org/stable/41471999. https://doi.org/10.1556/SocEc.26.2004.2-3.3

Dipela, M. & Mohapi, J. 2021. Barriers affecting effective monitoring and evaluation of poverty alleviation projects within Waterberg district. *Social Work / Maatskaplike Werk,* 57(3), pp. 287-301. doi:10.15270/522948

Goldman, I., Olaleye, W., Ntakumba, S. S., Makgaba, M. & Waller, C. 2021. *Mere compliance or learning - M&E culture in the public service of Benin, Uganda and South Africa.* In I. Goldman, & M. Pabari (Eds.), Using evidence in policy and practice: Lessons from Africa. New York: Routledge. https://doi.org/10.4324/9781003007043-4

Hatry, H. 2008. *Emerging developments in performance measurement: an international perspective.* In the International handbook of practice-based performance management. pp. 4-24. SAGE Publications Inc. doi:10.4135/9781412982719

Holvoet, N. & Renard, R. 2007. Monitoring and evaluation under the PRSP: Solid rock or quick sand? *Evaluation and Program Planning,* 30, pp. 66-81. https://doi.org/10.1016/j.evalprogplan.2006.09.002

Holvoet, N., Gildemyn, M. & Inberg, L. 2012. Taking stock of monitoring and evaluation arrangements in the context of Poverty Reduction Strategy Papers: Evidence from 20 aid-dependent countries in sub-Saharan Africa. *Development Policy Review*, 30(6), pp. 749-772. https://doi.org/10.1111/j.1467-7679.2012.00597.x

Huovila, A., Bosch, P. & Airaksinen, M. 2019. *Comparative analysis of standardized indicators for Smart sustainable cities: What indicators and standards to use and when?* Cities, 89, pp. 141-153. doi:10.1016/j.cities.2019.01.029.

Kawooya, I., Ddumba, I., Kayongo, E. & Mijumbi-Deve, R. 2021. *Rapidly responding to policy queries with evidence: Learning from rapid response services in Uganda.* In I. Goldman, & M. Pabari (Eds.), Using evidence in policy and practice: Lessons from Africa. New York: Routledge. https://doi.org/10.4324/9781003007043-8

Kusek, J. Z. & Rist, R. C. 2004. *Ten Steps to a Results-Based Monitoring and Evaluation System.* Washington DC: The World Bank.

Lahey, R. 2015. *Common isues affecting monitoring and evaluation of large ILO projects: Strategies to address them. i-eval THINK Piece, 9.* https://www.ilo.org/wcmsp5/groups/public/---ed_mas/---eval/documents/publication/wcms_424047.pdf

Lamhauge, N., Lanzi, E. & Agrawala, S. 2013. The use of indicators for monitoring and evaluation of adaptation: Lessons from development cooperation agencies. *Climate and Development*, 5(3), pp. 229-241. doi:10.1080/17565529.2013.801824

Lennie, J., Tacchi, J., Koirala, B., Willmore, M. & Skuse, A. 2011. *Equal Access Participatory Monitoring and Evaluation Toolkit*. Queensland University, Equal Access, University of Adelaide, USAID and Australian Government. https://www.betterevaluation.org/en/toolkits/equal_access_participatory_monitoring

Maleka, E. 2017. Monitoring and evaluation of sport-based HIV/AIDS awareness programmes: Strengthening outcome indicators. SAHARA-J: *Journal of Social Aspects of HIV/AIDS,* 14(1), pp. 1-21. doi:10.1080/17290376.2016.1266506

Mark, M. M. & Henry, G. T. 2004. The mechanisms and outcomes of evaluation influence. *Evaluation,* 10(1), pp. 35-57. doi:10.1177/1356389004042326

Moreno Pires, S., Magee, L. & Holden, M. 2017. Learning from community indicators movements: Towards a citizen-powered urban data revolution. *Environment and Planning C: Politics and Space,* 35(7), pp. 1304-1323. doi:10.1177/2399654417691512

Reed, M. S., Dougill, A. J. & Baker, T. R. 2008. Participatory indicator development: What can ecologists and local communities learn from each other? *Ecological Applications*, 18(5), pp. 1253-1269. https://doi.org/10.1890/07-0519.1

Reynolds, M. 2014. Equity-focused developmental evaluation using critical systems thinking. *Evaluation*, 20(1), pp. 75-95. doi:10.1177/1356389013516054

Sen, N., Kessler, A. & Loveridge, D. 2018. *Guidelines to the DCED standard for results measurement: Defining indicators of change and other information needs*. https://www.enterprise-development.org/wp-content/uploads/2_Implementation_Guidelines_Defining_Indicators.pdf

Silitonga, N., Ruddick, A., Widjajanti, P., Yatim, D., Blogg, J., Blogg, S. & Movono, I. 2013. Challenges of Monitoring and Evaluating an AusAID-funded HIV Program in Indonesia. *Evaluation Journal of Australasia,* 13(1), pp. 36-43. doi:10.1177/1035719X1301300106

UNAIDS. *An introduction to indicators*. https://www.unaids.org/sites/default/files/sub_landing/files/8_2-Intro-to-IndicatorsFMEF.pdf

USAID Resilient Waters. 2019. *Baseline report: Understanding resilience in the Limpopo and Okavango river basins*. https://reshubafrica.org/wp-content/uploads/2020/09/All-sites_USAID-Resilient-Waters-baseline-report-ENGLISH.pdf

Valadez, J. & Bamberger, M. 1994. *Monitoring and evaluating social programs in developing countries*. Washington DC: World Bank. https://doi.org/10.1596/0-8213-2989-8

Weiss, C. H., Murphy-Graham, E. & Birkeland, S. 2005. An alternate route to policy influence: How evaluations affect D.A.R.E. *American Journal of Journal of Evaluation*, 26(1), pp. 12-30. doi:10.1177/1098214004273337

World Bank. 2007. *Results of the expert roundtables on innovative performance measurement tools*: Note #3: Tools to improve the quality of indicators. Washington DC. http://web.worldbank.org/archive/website01404/WEB/IMAGES/TECHNI-4.PDF

Chapter 7: Monitoring in an Era of Algorithmic Governance

Rekgotsofetse Chikane and Halfdan Lynge

Introduction

In a world that has become more connected, interdependent, and data-rich, and where Internet users tripled from 1 billion in 2005 to 3.2 billion in 2015, digital services have increasingly allowed new actors to become producers, owners, and consumers of data (Bamberger 2016:29). The rapid increase in data and information across the various aspects of our society has altered the logic and way institutions and organisations operate internally and externally. However, as argued by York and Bamberger (2020), the current usages of big data and data analytics have focused on areas of implementation, coordination, the management of programmes, strategies, and services but has not found wide scale usage with regards to the monitoring and evaluation (M&E) of these efforts and programmes. For instance, the administrative data produced by public institutions alone, if mined appropriately, provides a plethora of opportunities to better understand "societal patterns, trends and policy impacts" and, if sanitised and released to the broader public, could fuel the innovation of products and services unbeknownst to these institutions themselves (Veale and Brass 2019:2).

Our point of departure is 'datafication' (Mayer-Schönberger and Cukier 2013), which refers to "the ability to transform non-traditional information sources such as text, images, and transactional records into data" (Diermeier 2015). Datafication has "created the opportunity for policymakers to have deeper, data-driven insights" (Nuamnovic 2017) and "allowed quantitative analysis to penetrate the policy process more deeply than ever before" (Diermeier 2015). As a result, as Denick et al. (2019:3) argue that it has become increasingly important to understand the technical ability to convert "increasing amounts of social activity and human behaviour into data that can be collected and analysed". This, in turn, demands an investigation of the relationship between M&E specialists and data scientists.

The use of data, particularly within the public sector, is established and has been of central concern for the M&E of government activities. Performance measurement places a central focus on the relationship between inputs, outputs, outcomes, benchmarks, citizen satisfaction, productivity, and the tools available to observe the data generated by these distinct but interdependent entities, and whether they achieve the public objectives (Williams 2003:643).

It also poses the question of the role of data science in understanding various public administration processes concerning the evolving nature of public governance caused by increased datafication and digitalisation. The reforms within this sector are driven by digital service transformations such as Industry 4.0 (I4.0), 'e-government 3.0', and the creation of integrated data infrastructures within the state, which were designed to improve the experience of citizens by making governments more efficient (Veale and Brass 2019:2). To achieve this, there is also a need to look into the operations and systems within government and the role that datafication and digitalisation play in the transformation of these structures.

A further need is to understand the potential impact that algorithmic governance will have on monitoring systems within the state. 'Algorithmic governance' here refers to a form of government of social ordering, where the usage of algorithms (especially artificial intelligence) is applied to public policy development and implementation. It does not refer to the governance of algorithms.

Considering these changes, what remain certain for evaluators and data scientists are three uses of data for policy maintenance and generation: i) that data and evidence should be used to effectively inform decisions on appropriate policy action; ii) data generated from the evaluation of policies should be used to inform decisions regarding whether to continue, halt, or improve policies; and iii) that evidence is used to inform the future consideration of policy options (Sanderson 2002:4).

A central concern for the use of data within the public sector, and monitoring systems in Africa in particular, is institutional readiness. This concern is in relation to institutional systems, cultures, and methodological approaches that are not in full alignment with approaches endorsed by data scientists. The barriers to the institutional preparedness of monitoring systems related to datafication in the public sector include the cost of data collection, the lack of IT infrastructure, and in other cases, the presence of various risks associated with legacy IT systems within the state. Other barriers include the lack of skilled data practitioners who can use data and understand the value of data generation, and the lack of governance frameworks focused on the management of data for the 21st century. This chapter works towards understanding these challenges as opportunities for both monitoring systems and data evaluators.

This chapter is premised on two assumptions about the public sector in Africa. The first is that it is highly politicised, meaning that politics influences decisions deep within civil service. There is nothing uniquely African about that. In the United States, the president and the cabinet secretaries can appoint over 4 000 employees, and countries previously thought to be immune to politicisation (e.g., Denmark and New Zealand) have succumbed to "creeping politicisation" (Halligan 2021:2). However, politicisation — political appointments throughout the civil service, partisanship in promotions, the use of ministerial advisors, the reliance on ministerial cabinets, etc. — is a common feature throughout the African continent, and something all African public servants must take into consideration.

The second assumption is that the African public sector suffers resource constraints with regards to finance and in terms of human resources. With low tax-to-GDP ratios, African governments have little fiscal space to attract and retain talent and invest in digitalisation and data collection. In 2019, the tax-to-GDP ratio in Africa was only 16.6%, according to a report by the Organisation for Economic Development and Cooperation, the African Union Commission, and the African Tax Administration Forum (2021). By comparison, it was 22.9% in Latin America and the Caribbean, and 33.8% among OECD countries.

The chapter is structured as follows: Section 1 discusses the drivers of datafication and their implication on data analytics. Section 2 details how datafication has impacted both M&E specialists and data scientists. Section 3 examines the methodological differences between the approaches to monitoring between M&E specialists and data scientists. Section 4 details how monitoring has changed as governance systems have evolved, culminating with the emergence of algorithmic governance[23] and the broad challenges it has created. Section 5 highlights the democratic and technocratic implications of algorithmic governance for both evaluators and data scientists.

Drivers of Datafication

The argument in this chapter is based on three observations. The first is that data volumes are increasing. Figure 1 plots the change in global data volumes from 2010 to 2025. It suggests that by 2025 global data volumes will reach 181 zettabytes, up from 64.2 zettabytes in 2020.[24] Since 2010, data volumes have doubled every two to three years. This means around one-fourth of the data ever produced was produced last year, while 80% was produced over the course of the last five years. Nothing suggests that the trend is about to reverse. According to the International

---

23 'Algorithmic governance' here refers to a form of government of social ordering, where the usage of algorithms, especially artificial intelligence, is applied to public policy development and implementation. It does not refer to the governance of algorithms.

24 One zettabyte is 10^{21} (10 followed by 21 zeros) bytes: approximately the amount of data that can be stored in 250 billion DVDs.

Telecommunication Union, 63% of the world's population is now online (2021). Every day, they send 300 billion emails (Lynkova 2021), 100 billion WhatsApp messages (Singh 2020), and 8 billion messages on Facebook (Bulao 2021). They post more than 500 million stories on Instagram (Chernev 2021) and upload 720,000 hours of videos on YouTube (Petrov 2021). An estimated 46 billion Internet of Things (IoT) devices also contribute to the increasing data volumes (Galov 2021). For example, the 770 million surveillance cameras currently installed produce more than nine quadrillion images every day (Bischoff 2021), assuming an average frame rate of 15 frames per second.
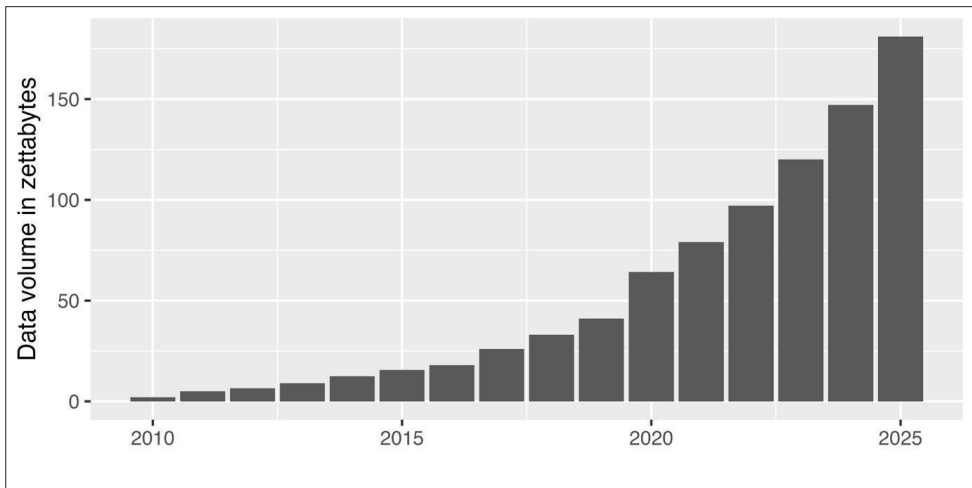

Figure 7.1: Reinsel et al., *Global Data Volumes*, 2010 - 2015

While a lot of data is produced, a lot less is stored. Rydning and Reinsel (2021) estimate that global storage capacity will only reach 16 zettabytes by 2025, which means that 90% of the data produced will be lost. Data is generally stored in three locations. The first is called endpoints, which are the personal computers, mobile phones, IoT devices, etc., that produce most of the data. The second location is the edge, including mobile phone towers, institutional servers, and smaller data centres set up to reduce response times and meet privacy requirements. The third and final location is the core, which includes traditional servers and cloud-computing data centres. There are around 600 hyperscale data centres (centres with more than 5 000 servers) in the world. The two largest are the China Telecom Data Centre in Hohhot, China, which occupies one square kilometre, and The Citadel in Tahoe Reno, the United States, which occupies 700 000 square metres and uses 815 megawatts of power. To meet the growing demand for storage capacity, around 50 hyperscale data centres are built every year.

The second observation made in this chapter is that the nature of data is changing. Data scientists make a distinction between 'structured' and 'unstructured' data. Structured data refers to data that is neatly organised in rows and columns, such as an Excel spreadsheet detailing information from a survey. Each row typically represents a respondent, while each column captures information about that respondent: gender, age, level of education, employment status, income, etc. The spreadsheet may miss information or contain empty cells, but it does not fundamentally alter the structure of the data. Unstructured data, by contrast, is unorganised. Rather, it is not organised in any predefined manner. Examples of this are text, images, videos, sound, etc. Reinsel et al. (2021) predicts that 80% of global data volumes will be unstructured by 2025.

A lot of the unstructured data is 'data exhaust', which is what the trail of data we leave behind when we browse the Internet is called. Every click, scroll, or hovering of a cursor produces data stored in the form of cookies, temporary files, logfiles, storable choices, etc. The data is captured to improve individual online experiences through the customisation of content. However, it can also be used to generate knowledge. Google, for example, uses data exhaust to optimise the placement of advertisements (Zuboff 2015). A company's advertising services are offered under a pay-per-click pricing model, which results in advertisers exclusively paying under the condition of a Google user clicking on their advertisement. To maximise profits, Google uses cookies and keywords determined by advertisers to match advertisements with Google users who may be interested and are more likely to divert their browsing. This matching of advertisements and Google users is at the core of Google's business model. In 2020, Google Ads, the company's advertising platform, was the main source of revenue for Alphabet Inc., Google's parent company, contributing USD 168.6 billion (Alphabet Inc. 2021).

The changing nature of data also means the boundary between research methods is becoming increasingly blurred. Scientists make a fundamental distinction between quantitative and qualitative data, where the former refer to numeric data, while the latter refers to non-numeric data. The distinction has shaped our thinking around data and research methods. Most social science degrees and programmes, for example, separate their quantitative and qualitative methods training, and many social scientists identify as either quantitative or qualitative researchers. The distinction has also contributed to the 'paradigm war' (Gage 1989). This was originally a debate between different ontological and epistemological positions in many fields, and it developed into a debate over the merits and assumptions of different research methods (Bryman 2008). However, recent advancements in data science are beginning to erase the boundary between the two methods. All types of data — whether numeric or non-numeric — can now be processed through the use of a computer.

Text is perhaps the most typical example of qualitative data, and documents, transcripts, field notes, etc. are important sources of information for qualitative researchers. However, text can also be converted into numbers. This is done in social media post analysis processes. In theory, social media posts are qualitative or non-numeric data. However, since they are published online, in practice, they are not. Text mining is an AI technology that uses natural language processing (NLP) to transform unstructured text into structured data, suitable for analysis or to drive machine learning (ML) algorithms. It can identify facts, relationships, and patterns that would otherwise remain hidden for the qualitative researcher, buried in the mass of textual big data.

Our third and final observation is that processing power is improving. In 1975, Gordon Moore, the co-founder of Intel, one of the world's largest semiconductor manufacturers, predicted that the number of transistors in an integrated circuit would double every two years. Moore's prediction, which has since become known as Moore's Law, has largely proven true. Figure 2 plots the number of transistors in an integrated circuit from 1971 to 2017. In 1971, this number was 2 308. By 2017, it had increased to 19.2 billion, giving an annual growth rate of around 42%: just short of Moor's prediction.
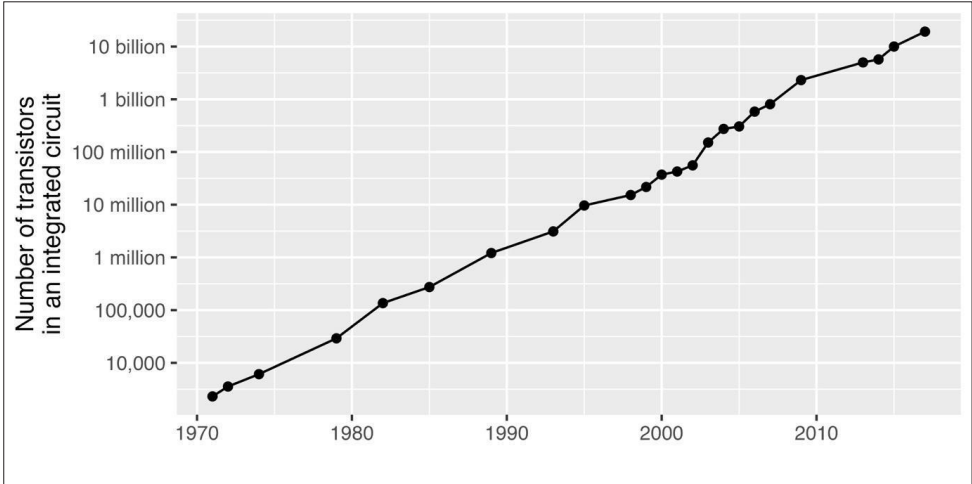
Figure 7.2:   Rupp, *Transistors in an integrated circuit*, 2018

The importance of the number of transistors in an integrated circuit or a microprocessor is due to it being the determining factor in the processing power of a computer. Transistors in a microprocessor can be compared to cylinders in a car engine: the more cylinders, the more horsepower, the faster the car drives. Similarly, the more transistors in a microprocessor, the more processing power, and the faster a computer can execute commands.

Moore's Law has pushed the boundaries of data science and enabled us to work with larger datasets and techniques that previously took days or weeks to implement or existed only in theory. Machine learning (ML) can be taken as an example of this. In 1959, when Arthur Samuel developed his checkers-playing programme (Samuel 1959), which was the first successful implementation of ML, he used an IBM 704: a computer the size of 15-20 large refrigerators, which could execute around 40 000 commands per second. Today, far more complex algorithms can be run even on ordinary laptops, and ML has become a standard tool for most data scientists. In fact, advances in ML are driving an entirely new type of 'autonomous analytics'. Historically, analytics were for decision-makers who were expected to consider the outputs and make decisions. With advancements in ML, computers can do this in their stead. This means that corrective action can be taken much faster, an example being instances of policies having unexpected negative spillover effects. However, it also raises questions of accountability. This is further discussed below.

Bringing M&E Specialists and Data Scientists Closer Together

Monitoring activities exist within an interventionist culture that focuses on evidence-informed decision-making. As Head (2008:2) explains, the state's use of evidence is ideally linked to its aspiration to produce knowledge that would be required for the "fine-tuning of programs and constructing guidelines and 'tool-kits' for dealing with known problems". However, various challenges persist regarding the integration of M&E specialists and data scientists relating to differences in institutional cultures and methodological approaches. These differences explain how M&E specialists and data scientists utilise different research and management paradigms consisting of different terminologies and approaches to similar questions, i.e. how to monitor interventions, the nature, quality, and utility of various types of data, and the role of theory informing data usage (Bamberger 2016).

An analytical framework devised by Langer and Weyrauch (2020) provides a valuable tool to understand the culture that M&E specialists find themselves situated within. For Langer and Weyrauch, three factors are important. The first is the demand for evidence within the state which relates to the generation, use thereof, and change mechanisms related to evidence. The second relates to the external context that informs the relationship between the state, its macro-political context of evaluation, and its various stakeholders. Lastly, the internal context of a state's M&E processes is influenced by culture, organisational capacity, management, and resources at the state's disposal (Langer and Weyrauch 2020). The changes in governance due to datafication will inevitably alter how M&E specialists engage with all three factors.

To further augment this framework, the nature of knowledge that M&E specialists must consider is also of importance. Head (2008) provides a useful intervention noting that evidence-informed decisions must contend with the political, scientific,

and practical implementation of knowledge. The policy relates to the external political constraints, the limits of data, and the nature of interventions employed (Head 2008). Scientific implementation of monitoring and evaluation relates to the chosen methodologies used by M&E specialists to gather and analyse information. Two broad fields inform the methodological use of data: obtaining methodological rigour through experimentation with data for evidence-informed decision-making, or a more hermeneutic approach that emphasises iterative learning based on research projects (Head 2008). The practical implementation of knowledge refers to the ability for evaluators to understand the internal institutional challenges to the implementation of programmes that require a level of management that is upwards, downwards, and outwards with internal stakeholders (Head 2008).

The developments within the field of data science have created a variety of challenges relating to the external context that monitoring services that exist within the state, the practical utilisation of data for monitoring, and the actual strategies of using data that confront traditional means of monitoring the public sector. Figure 3 illustrates how both frameworks are combined to provide a means to compare both M&E specialists and data scientists within the context of increased datafication. Table 1 compares how the framework relates to both M&E specialists and data scientists. We use the framework to provide a means of comparing how datafication impacts both M&E specialists and data scientists within monitoring systems created by the state.
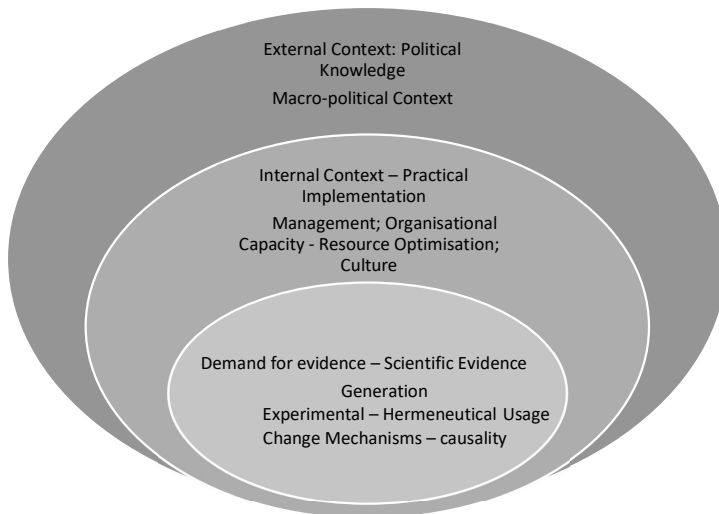

Figure 7.3: Head and Weyrauch, *Framework combination and comparison*, 2008 and 2020

Table 7.1:   M&E Specialists and data scientists

	Evaluators	Data Scientists
External Context: Political Knowledge	Macro-Context Intra-relationships with stakeholders and non-state agents	Politics of Data generation and usage Privacy Surveillance Legal and regulatory requirements
Internal Context: Practical Implementation	Culture Organisational Capacity Management Resource Optimisation Individual/Organisational/Systems Change Motivation to use evidence Capability to use evidence Opportunity to use evidence	Data Governance Governance Mechanisms Structural Mechanisms Procedural Mechanisms Relational Mechanisms
Demand for Data – Scientific Data usage	Evidence Generation Use Intervention Change Mechanisms	Data Scope Domain Scope

How Datafication Augments M&E at the External Level

The drivers of change that inform the evolving nature of data and governance have various implications for traditional M&E specialists and data scientists within the public sector. Both must contend with external political constraints placed on their activities. As seen in Table 1 and Figure 3, the external cultural context described by Langer and Weyrauch relates to Head's considerations of the political constraints placed on the use of data. While M&E specialists focus on the broader political context of M&E and its uses by the state, data scientists within the public sector must contend with the increasing political suspicion regarding the use of data by public institutions. Datafication and its various usages has created a need for new legal and regulatory requirements for both M&E specialists and data scientists.

Increased levels of data generation by both the private sector and the state have resulted in a resurgence of interest by citizens, businesses, and non-state actors on issues such as privacy, surveillance, anonymity, and digital discrimination. This increased scrutiny has led to an increase in guidelines, regulations, and legislation related to data protection and the rights allocated to data users and generators (Ruppert et al. 2017), all of which inform the strategies undertaken by M&E specialists

and data scientists. The politicisation of data is less about data itself, but rather the political constraints placed on the use of data and its generation (Ruppert et al. 2017).

The outcome of this renewed interest in governments' use of data to inform policy decisions results in the belief that individuals need to protect themselves from ensuing politics related to data generation. The cause of this is varied, though according to Hilden (2019), it results from the increased marketisation of personal information continuously generated by individuals. Though data remains value-neutral, data itself provides varying levels of economic value to various actors in society. Increased marketisation of personal information results from the understanding that "data on individuals and their performance generates greater efficiency and control" over social behaviours that can be quantified, and individual behaviour can be predicted (Hildén 2019:29).

How Datafication Augments M&E at the Internal Level

The internal institutional context of evaluators is informed by practical limitations of using data. As seen in Table 1, for M&E specialists, these limitations relate to organisational capacity, resources, and the management of data amongst others. All of these inform the internal institutional context of monitoring, how data scientists must contend with the mechanisms related to the structure of data, the procedures related to its usage, and how data within an institution relates to its various functions. Improved computational possibilities have naturally created new management information systems that will inevitably be integrated into programme designs, management, monitoring activities, and evaluation studies (Ruppert et al. 2017:200; York and Bamberger 2020).

Data governance within the state becomes an essential consideration for any monitoring activity. The use of data analytics in the public service primarily seeks to categorise, segment, rate, and rank specific populations, administrative activities, and various transactional datasets. Once data is appropriately organised, the intent of data analytics and data governance becomes the efficient allocation of resources and services.

According to Abraham, Schneider, and Brocke (2019:425), based on a systematic review of literature of data governance, "data governance specifies a cross-func-tional framework for managing data as a strategic enterprise asset. In doing so, data governance specifies decision, rights, and accountabilities for an organisation's decision-making about its data. Furthermore, data governance formalises data policies, standards, and procedures, and monitors compliance".

Abraham et al. (2019) argues for an understanding of data governance through its relation to governance mechanisms that comprise formalised structures that connect the activities of the state with its IT capabilities and data management

functions. This includes the "formal processes and procedures for decision-making and monitoring, and practices supporting the active participation of and collaboration among stakeholders" (Abraham et al. 2019:428). These mechanisms are structural, procedural, and relational. Structural mechanisms refer to reporting structures that indicate the roles and responsibilities of data and governance bodies that allocate decision-making authority regarding the use of data. Procedural and relational mechanisms refer to how "data is recorded accurately, held securely, used effectively, and shared appropriately", and the management of the collaboration of data amongst stakeholders (Abraham et al. 2019:429).

How Datafication Augments the Methodological Approaches to Monitoring

Datafication and advancements towards algorithmic governance will undoubtedly force M&E into a new realm where traditional approaches and techniques, though helpful, may not be the most efficient means of creating an understanding of public sector performance and impact. As an example, M&E specialists may soon be required to generate or synthesise new technologies based on existing data, including data exhaust, rather than collect new data (York and Bamberger 2020). Data scientists already have various tools at their disposal to account for this across the policy and programme cycle.

Rather than relying on the diagnostic, design, implementation, and outcome or impact cycle of policy development, data scientists view data that informs policy through the framing of descriptive and exploratory analysis, predictive analytics, detection, and evaluation or prescription (Bamberger 2016). The descriptive and exploratory analysis focuses on the compilation of large datasets that would ordinarily be beyond the capacity of conventional tools. For instance, the use of real-time data collection tools (for both structures and unstructured datasets) allows for the dynamic monitoring and generation of actionable data on project problems as well as new opportunities. Evaluators rely on more familiar tools such as surveys, project records, and official statistics (Bamberger 2016; York and Bamberger 2020:25). Furthermore, while evaluators may utilise theory-based approaches to evaluation design, the use of theory is less specific for data scientists (Bamberger 2016; Langer and Weyrauch 2020).

For M&E specialists, experimental or quasi-experimental research focuses on assessment of the causal relationships between an intervention and its intended outcomes by controlling for other factors that could affect the outcomes (York and Bamberger 2020). It is an approach taken due to the belief that drawing upon the correlation between variables does not provide practical utility for policymakers (York and Bamberger 2020). By contrast, data scientists utilise various techniques on constantly updating large datasets which cover a wide range of variables to create predictive means of assessing how certain groups are most likely to respond

to specific interventions (York and Bamberger 2020). This approach allows data scientists to detect real-time relationships that conventional M&E tools could not or were not designed to detect.

The ability to utilise predictive analytics is due to large datasets being able to adequately represent the total population size. Though challenges exist with this assertion, such as selection bias, data quality, and the proprietary algorithms used to generate this data, improvements in this sector seek to address these concerns (York and Bamberger 2020). Increased datafication allows for techniques such as data mining, computational analysis, and predictive modelling on large datasets to make informative and reliable predictions.

M&E specialists and data scientists also potentially clash regarding epistemology. Public policy is generally theory-driven, and monitoring is restricted to a predefined set of indicators. With data science and the increase in data volumes and processing power, we can expand this set of indicators and monitor spillover effects outside of the target sector or population. Conditional cash transfer, education, and health programmes often target the poor. However, the poor are often a subset of a local economy, loosely defined as the geographical unit within which the poor live, such as the city, the municipality, the neighbourhood, or the household.

Programmes may also affect non-target populations. For example, conditional cash transfer recipients may purchase goods and services from ineligible households, thus contributing to their livelihoods. Similarly, children who receive free textbooks and computers may share them with other children (e.g., relatives and friends), thereby increasing their enrolment. Finally, supplying deworming drugs to the most exposed children may also benefit the less exposed children by reducing disease transmission and lowering infection rates. Sometimes these spillover effects are predicted. Often, they are not, which means they are unlikely to be detected using traditional, deductive monitoring. Inductive monitoring presents its own methodological challenges. While we can detect changes, we generally cannot attribute changes to a specific programme. In other words, we may have suggestive evidence that project X had spillover effects on target sector or population Y. However, we cannot prove that X caused Y.

Towards Algorithmic Governance

Governance broadly relates to the coordination of social interactions through the production and implementation of ideas, plans, regulations, and policies concerning the public to solve collective action problems (Gritsenko and Wood 2020). It is argued that the increase in data volumes, the changing nature of data, and the improvements in processing power have implications for governance in general, and monitoring. As public institutions embrace the opportunities brought by datafication, new challenges emerge against established means of governance.

The evolution of governance during both the 20th and early 21st centuries has been predicated on the need for the state to produce new solutions to problems created by changing social and economic conditions within society. Though the form and focus of these reforms in the public sector differ worldwide, the reforms remain a shared experience and are often seen as a means to multiple ends (Pollitt and Bouckaert 2017). Whether these ends relate to creating savings in public expenditure, improving the quality of public service, creating more efficient government operational and administrative systems, or improving the effectiveness of government interventions, the need for change is imperative (Pollitt and Bouckaert 2017). Even though M&E has had its own evolutionary path as a discipline in its own respect, the rearrangement of public institutions and administrative processes over time has influenced the usage of M&E tools at the public servants' disposal. What remains consistent is that decision-makers who build programmes intended to create social reforms rely on relevant and usable knowledge (Head 2008). However, as the nature of relevant and usable knowledge experiences exponentially increases, it begs the question of how governance should alter itself to accommodate these changes.

Industry 4.0 (I4.0) refers to the integration of information and operational technologies in a manner that transforms the production processes and "takes advantage of the abundant information available in each stage of the value chain, from suppliers to customers, of any industrial sector (manufacturing, energy, transport, supplies, mining, health, pharmaceutical, etc.)" (Zorrilla and Yebenes 2022:1). I4.0 can be expanded towards the public sector. The impact of the transformations caused by I4.0 has necessitated the integration of a ubiquitous level of connectivity between data, people, processes, and services, where all of these elements exchange and exploit each other's information, leading to further increase in both the amount and variation of data generated in real-time (Zorrilla and Yebenes 2022). These advancements in digital technologies and datafication have ensured that there has been an increased level of "information-based steering of activities", often with the intent of creating greater levels of coordination when trying to solve complex societal problems (König 2020:469). These activities refer to the generation, transfer, storage, and processing of information through networked interactions characterised by the "co-presence of manifold entities which can dynamically adjust their behaviours" (König 2020:469).

Datafication and the increased networks of stakeholders working on societal issues have necessitated a new form of governance: algorithmic governance. Algorithmic governance refers to the usage of fine-grained information generated from various entities that allow for the finding of patterns in the behaviours and interactions of these entities with the final intent of coordinating these behaviours in the future (König 2020). It refers to the organisation of society and the economy by reiterating step-by-step processes and/or rules that determine how inputs are processed into outputs. It results in a form of governance less focused on administrative processes, but rather the governance of public services using knowledge about "service users

and citizens that is collected from them in order to govern them more effectively" (Williamson 2014:3). Governance then becomes a set of "processes of decentralised coordination of distributed entities" whose changing behaviour feeds back into the algorithmic decision-making system so that the behaviours of the other entities or agents can be attuned to these information updates (König 2020).

The shift to algorithmic governance signifies the intent by the state to contend with developments within society ranging from the analysis of "big data' generated from transactional processes, peer production, the democratisation of innovation, crowdsourcing, wikinomics, cognitive surplus, and network effects" (Margetts and Dunleavy 2013). Data in and of itself is value neutral. Nevertheless, as an input for algorithms, it can have both intended and unintended consequences due to data only providing a partial view of the world. As a result, its generation and interpretation may be biased or incomplete, or might be deliberately or accidentally manipulated before ever being used (Janssen & Kuk 2016). Regardless, the resultant shift to algorithmic governance has created the ability to design governance structures that utilise data about human-human and human-state interactions to create tailored policy interventions or predictions that meet the desires and interests of citizens in real-time (Gritsenko and Wood 2020).

In the past, these designs were limited to the dominant paradigm of the time. With algorithmic governance, these designs can be made to be hierarchical (the traditional model), self-governed (new public management through increased disaggregation), or co-governed (networked governance), depending on what the state requires. This is due to algorithmic governance allowing for the choice of governance architecture that can structure the decision situations of individuals within the public sector by providing "certain information, options, and suggestions, thereby making some choices more and others less likely" (Yeung 2017).

Three Broad Challenges

Algorithmic governance poses three broad challenges for monitoring (Gritsenko and Wood 2020). The first is automated decision-making. In traditional monitoring systems, decisions are taken by individuals who can be held accountable. Institutional arrangements inform these accountability structures at both a political and institutional level. By contrast, in automated decision-making systems, accountability is subverted, and monitoring is opened to various challenges such as automation bias, the deprivation of humans as agents of judgement and reason, and the development of self-perpetuating echo chambers (Gritsenko and Wood 2020). Furthermore, automated decision-making ensures that the human element no longer functions within the traditional causal linkages provided by statistical analysis. The M&E specialist no longer plays a central controlling role regarding what variables are "compared and combined to generate predictive outputs" (Appel and Coglianese 2020:166).

The second broad challenge is the invisibility of decision-making. This poses a renewed challenge on monitoring by providing a means of decision-making that may usurp public interest (Gritsenko and Wood 2020). Though a policy's political mandate may be unambiguous, the circumvention of public interest through automated decision-making based on data generated by citizens, businesses, or the state entails altering a policy in a manner that may deviate from its original intent. Without traditional structures of accountability, this alteration creates a "black box society" where the changes to policies are not reliant on changes in public interest, but rather the changes observed by what may seem to be ambiguous algorithmic determinations of what is, or may be, of interest to the public. As Appel and Coglianese (2020:167) explain, the black box of algorithmic governance is the result of the inherent difficulty in explaining how the outputs are generated to the public and the inability to ascribe a "causal relationship between an algorithm's input data and its output prediction".

The third broad challenge relates to ethics. Though data may be politically neutral, the choices and base assumptions that inform algorithmic processes are informed by the designers of these algorithms, creating a wide variety of intended and unintended consequences that pose specific ethical questions (Gritsenko and Wood 2020). Though automated decisions based on algorithms may be happily accepted when creating efficiencies within a postal office, it is unclear how these decisions would be accepted regarding decisions related to the criminal justice system (Appel and Coglianese 2020).

Democratic Governance or Technocratic Solutionism

The three broad challenges posed above, brought about by datafication and the emergence of algorithmic governance, bring an old debate to the forefront which involves the balance between democratic governance and technocratic solutionism. M&E specialists, at both the policy implementation and the data generation level of analysis, must contend with the need to develop tools to assess and monitor technical solutions related to algorithmic interventions within the state while also assessing the democratic efficacies of these interventions. These solutions relate to automated decision-making interventions and predictive analytics that utilise data from "official records and statistics; secondary data obtained through administrative operations from frontend services; user-generated data often in the form of web content such as blogs, chats, tweets and videos; sensory data gathered by connected people and devices; tracking data such as CCTV, GPS, or traffic data, as well as satellite and aerial imagery; and transaction data from shopping or banking records" (Engin and Treleaven 2019:450). It also extends to solutions centred on the examination of big data to discover hidden patterns and undiscovered correlations within government interventions (Engin and Treleaven 2019).

At an institutional level, assessing the technical efficacies of algorithmic governance will increasingly fall upon the shoulders of M&E specialists and data scientists who will need to work in tandem to devise appropriate tools to monitor and assess algorithmic programmes. Though room will still exist for traditional approaches to governance, increased automation in decision-making, rapid increases in data generation and processing capabilities, and the predictive power of algorithmic approaches to monitoring will demand a new set of skills from M&E specialists. Skills will also require the support of reimagined governance mechanisms that better understand the relationship between human evaluators and computational code.

From a democratic perspective, algorithms remain inherently relational and reliant on human intuition. It involves balancing an intricate and dynamic arrangement of people and code where M&E specialists and data scientists work on the design and implementation of algorithmic logic (Janssen and Kuk 2016). As a result, the role of M&E specialists will be fixed in the macro-political context of monitoring and evaluating. Creating democratic accountability measures for algorithms depends on various social and political contexts and relies on certain systemic factors such as political will, effective legal institutions, and the rule of law (Basu et al. 2021).

From this democratic perspective, the monitoring of algorithmic interventions relies on two general focus areas. The first relates to monitoring the adoption and implementation of algorithmic interventions and their consistency with the existing legal framework. M&E specialists within this space will be responsible for monitoring how algorithmic interventions remain aligned with an existing mechanism of administrative accountability. However, noting the rate at which the sector is evolving and the ambiguity of the potential scope of these interventions, it bears noting that specific interventions may not yet have clear legislative guidelines. Therefore, evaluators must be responsible for the identification of algorithmic interventions that deviate from democratic principles enshrined by overarching forms of legislation, such as constitutional frameworks and the various degrees of scrutiny applied to various algorithmic systems (Basu et al. 2021).

The second general focus area relates to transparency. The transparency of algorithmic systems will become dependent on the reporting done by monitoring systems specifically designed for these interventions. Transparency relates to the "specific decisions made about particular individuals or groups, as well as more general concerns around how the use of algorithmic systems is contributing to the function of particular public agencies, including policy-making and administrative functions" (Basu et al. 2021:45).

Due to the nascent nature of algorithmic governance within the continent, providing details of uses cases is difficult, and as a result, literature that provides deeper insights into its implication is scarce. That said, one case can be briefly discussed. This case involves Nigeria's implementation of its Rapid Response Register (RRR), which forms part of the federal government's Economic Sustainability Plan and the

National Social Protection Policy (Adeyeye 2022:11). The RRR functions as a means to scale the National Social Safety-Nets Programme and improve the standard of living of 100 million Nigerians identified as living in poverty (Apera et al. 2021:36). Utilising a combination of remote sensing, machine learning, and big data analysis that improves upon the traditional data sources previously used by the state, the RRR ranks urban poor wards according to their patterns of wealth and poverty (Apera et al. 2021:39). The use of these technologies provides the means for geographic targeting and poverty mapping that when used in conjunction with SMS and USSD application processes linked to mobile phone numbers, and the system provides the means for communities to self-register and enrol as beneficiaries of the RRR (Apera, Daniel, Balogun et al. 2021:40). The system improves the identification strategies of households eligible for government funded cash transfers, with validation of these households undertaken through SMS-based applications (Lowe, McCord and Beazley 2021:21). The system differs from traditional approaches to monitoring cash transfer systems as it is less reliant on community inputs collected by surveyors, but functions through the assessment of previously collected data validated through geographic targeting of citizen responses (Lowe et al. 2021:28).

The combination of these various technologies led to an overall improvement of cash transfers with 95% of identified eligible participants being both registered on the system and provided with a cash transfer (Apera et al. 2021:41). The targeting approach undertaken by the RRR was the first of its kind within sub-Saharan Africa and will be used to further scale Nigeria's other social protection initiatives and interventions (Apera et al. 2021:41). It further showcases the use of automated decision-making as a replacement for traditional methods of monitoring household wealth within urban poor regions in West Africa.

Conclusion

Datafication is changing our societies. It has already changed how public institutions operate, including the way in which policies are developed and implemented, and how public services are delivered. However, it has not yet fundamentally affected the monitoring of these policies. The potential is enormous. The utilisation of big data to monitor the performance of different policies and public institutions is one of the most exciting promises of the 'data age'. This chapter has discussed the opportunities and challenges associated with greater data utilisation in public sector monitoring. Three drivers of datafication were highlighted — the increase in data volumes, the changing nature of data, and the improving processing power of computers — which fundamentally change the context within which monitoring occurs. The realignment will require closer collaboration between M&E specialists and data scientists, who work within different institutional cultures and face epistemology and research methods in fundamentally different ways. Bringing them together will be an important challenge to overcome in the era of algorithmic governance.

Reference List

Abraham, R., Schneider, J. & vom Brocke, J. 2019. Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, *49*, pp. 424–438. https://doi.org/10.1016/j.ijinfomgt.2019.07.008

Adeyeye, A. 2022. *Understanding Public Policy and Good Governance in Nigeria* pp. 1–13. https://doi.org/10.1007/978-3-319-31816-5_4058-1

Alphabet Inc. 2021. *Alphabet Announces Fourth Quarter and Fiscal Year 2020 Results.* https://abc.xyz/investor/static/pdf/2020Q4_alphabet_earnings_release.pdf?cache=9e991fd

Apera, I., Daniel, A., Balogun, M., Mohammed, M., Lawal, A., Abdullahi, K. & Nongo, S. 2021. National Social Register: An Approach to Social Protection and Poverty Reduction in Nigeria. *Radical Statistics Newsletter*, *129*, pp. 29–44.

Appel, S. M. & Coglianese, C. 2020. Algorithmic Governance and Administrative Law. In W. Barfield (Ed.), *The Cambridge Handbook of the Law of Algorithms,* pp. 162–181. Cambridge University Press. https://doi.org/10.1017/9781108680844.009

Bamberger, M. 2016. Integrating big data into the monitoring and evaluation of development programmes. *New York: Global Pulse/Rockefeller Foundation.*

Basu, T., Brennan, J., Kak, A. & Joshi, D. 2021. *Algorithmic Accountability for the Public Sector.* Ada Lovelace Institute (Ada), AI Now Institute (AI Now), and Open Government Partnership (OGP). https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/

Bischoff, P. 2021. *Surveillance Camera Statistics: Which City has the Most CCTV Cameras?* https://www.comparitech.com/vpn-privacy/the-worlds-most-surveilled-cities/

Bryman, A. 2008. The End of the Paradigm Wars? In P. Alasuutari, L. Bickman, & J. Brannen, *The SAGE Handbook of Social Research Methods*, pp. 12–25. SAGE Publications Ltd. https://doi.org/10.4135/9781446212165.n2

Bulao, J. 2021. *30 of The Most Incredible Facebook Messenger Statistics for 2021.* TechJury. https://techjury.net/blog/facebook-messenger-statistics/

Chernev, B. 2021. *33+ Incredible Instagram Stories Statistics You Need to Know in 2021.* TechJury. https://techjury.net/blog/instagram-stories-statistics/

Dencik, L., Redden, J., Hintz, A. & Warne, H. 2019. The 'golden view': Data-driven governance in the scoring society. *Internet Policy Review*, *8*(2). https://policyreview.info/articles/analysis/golden-view-data-driven-governance-scoring-society. https://doi.org/10.14763/2019.2.1413

Diermeier, D. 2015. *Data science meets public policy.* The University of Chicago Magazine. https://mag.uchicago.edu/university-news/data-science-meets-public-policy

Engin, Z. & Treleaven, P. 2019. Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies. *The Computer Journal*, *62*(3), pp. 448–460. https://doi.org/10.1093/comjnl/bxy082

Gage, N. L. 1989. The paradigm wars and their aftermath a "historical" sketch of research on teaching since 1989. *Educational Researcher*, *18*(7), pp. 4–10. https://doi.org/10.3102/0013189X018007004

Galov, N. 2021. *How Many IoT Devices Are There in 2021? [All You Need To Know].* TechJury. https://techjury.net/blog/how-many-iot-devices-are-there/

Gritsenko, D., & Wood, M. 2020. Algorithmic governance: A modes of governance approach. *Regulation & Governance*, *n/a*(n/a). https://doi.org/10.1111/rego.12367

Head, B. W. 2008. Three Lenses of Evidence-Based Policy. *Australian Journal of Public Administration*, *67*(1), pp. 1–11. https://doi.org/10.1111/j.1467-8500.2007.00564.x

Hildén, J. 2019. The Politics of Datafication: The influence of lobbyists on the EU's data protection reform and its consequences for the legitimacy of the General Data Protection Regulation. *Publications of the Faculty of Social Sciences.*

International Telecommunication Union. 2021. *Individuals using the Internet*. ITU. https://www.itu.int:443/en/ITU-D/Statistics/Pages/stat/default.aspx

Janssen, M. & Kuk, G. 2016. The challenges and limits of big data algorithms in technocratic governance. In *Government Information Quarterly*, Vol. 33, Issue 3, pp. 371–377. Elsevier. https://doi.org/10.1016/j.giq.2016.08.011

König, P. D. 2020. Dissecting the Algorithmic Leviathan: On the Socio-Political Anatomy of Algorithmic Governance. *Philosophy & Technology*, *33*(3), 467–485. https://doi.org/10.1007/s13347-019-00363-w

Langer, L. & Weyrauch, V. 2020. Using evidence in Africa: A framework to assess what works, how and why. In *Using Evidence in Policy and Practice*, pp. 34–53. Routledge. https://doi.org/10.4324/9781003007043-3

Lowe, C., McCord, A. & Beazley, R. 2021. *National cash transfer responses to Covid-19: Operational lessons learned for social protection system-strengthening and future shocks*. ODI: Think Change. https://odi.org/en/publications/national-cash-transfer-responses-to-covid-19-operational-lessons-learned-for-social-protection-system-strengthening-and-future-shocks/

Lynkova, D. 2021. *The Surprising Reality of How Many Emails Are Sent Per Day in 2021*. TechJury. https://techjury.net/blog/how-many-emails-are-sent-per-day/

Margetts, H. & Dunleavy, P. 2013. The second wave of digital-era governance: A quasi-paradigm for government on the Web. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *371*(1987), 20120382. https://doi.org/10.1098/rsta.2012.0382

Mayer-Schönberger, V. & Cukier, K. 2013. *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.

Petrov, C. 2021. *YouTube Statistics You Will Need in 2021*. TechJury. https://techjury.net/blog/youtube-statistics/

Pollitt, C. & Bouckaert, G. 2017. *Public management reform: A comparative analysis-into the age of austerity*. Oxford university press.

Reinsel, D.,& Rydning, J. 2021. *Worldwide Global StorageSphere Forecast, 2021–2025: To Save or Not to Save Data, That Is the Question*. International Data Corporation. https://www.idc.com/getdoc.jsp?containerId=US47509621

Reinsel, D., Rydning, J. & Gantz, J. 2021. *Worldwide Global DataSphere Forecast, 2021–2025: The World Keeps Creating More Data — Now, What Do We Do with It All?* Internatiopnal Data Corporation. https://www.idc.com/getdoc.jsp?containerId=US46410421

Ruppert, E., Isin, E. & Bigo, D. 2017. Data politics. *Big Data & Society*, *4*(2), 2053951717717749.

Samuel, A. L. 1959. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, *3*(3), pp. 210–229. https://doi.org/10.1147/rd.33.0210

Sanderson, I. 2002. Evaluation, Policy Learning and Evidence-Based Policy Making. *Public Administration*, *80*(1), 1–22. https://doi.org/10.1111/1467-9299.00292

Veale, M. & Brass, I. 2019. *Administration by Algorithm? Public Management Meets Public Sector Machine Learning* (SSRN Scholarly Paper ID 3375391). Social Science Research Network. https://papers.ssrn.com/abstract=3375391. https://doi.org/10.31235/osf.io/mwhnb

Williams, D. W. 2003. Measuring government in the early twentieth century. *Public Administration Review*, *63*(6), pp. 643-659. https://doi.org/10.1111/1540-6210.00329

Williamson, B. 2014. Knowing public services: Cross-sector intermediaries and algorithmic governance in public sector reform. *Public Policy and Administration*, *29*(4), pp. 292–312. https://doi.org/10.1177/0952076714529139

Yeung, K. 2017. 'Hypernudge': Big Data as a mode of regulation by design. *Information, Communication & Society*, *20*(1), pp. 118–136. https://doi.org/10.1080/1369118X.2016.1186713

York, P. & Bamberger, M. 2020. Measuring results and impact in the age of big data: The nexus of evaluation, analytics, and digital technology. *The Rockefeller Foundation*.

Zorrilla, M. & Yebenes, J. 2022. A reference framework for the implementation of data governance systems for industry 4.0. *Computer Standards & Interfaces*, *81*, 103595. https://doi.org/10.1016/j.csi.2021.103595

Section 3

EXAMPLES OF MONITORING SYSTEMS IN PRACTICE

Chapter 8: Health and Demographic Surveillance Sites: Scalar and Spatial Monitoring

David Everatt

Introduction

This chapter analyses the challenges of monitoring via a health and demographic surveillance site (HDSS) in the heart of Gauteng, called GRT-INSPIRED. The GRT is a partnership between three major African universities, namely the University of Johannesburg, the University of Pretoria, and the University of the Witwatersrand. This HDSS node is split across three sites in Gauteng. The HDSS performs 'monitoring of a special type' which is detached from any specific programmatic intervention, focusing on one site with multiple visits each year. This monitoring system creates a data set that operates from the household through dwelling to the entire area under study, usually a population of around 100 000 people. The scale is important, as HDSS tracks individual movements in and out of their node and works with the entire nodal population, with various cohorts available for smaller studies, such as the elderly, women, or youth.

HDSS provides spatial monitoring because it is a site where every year, the people living in the node provide data on their health, demographics, socioeconomic status, migration, and the like. The HDSS measures this 'space' year after year to provide an accurate population-based form of monitoring. Data errors can be corrected because of the continuous fieldwork, and the HDSS data is a unique resource for planners and policy makers. The HDSS is also an evaluation. This is not an explicit categorisation, but an example of this is the continual measurement of the health status of poor people, which makes the HDSS an evaluation of the policy and service provision in its site.

In South Africa, the South African Population Research Infrastructure Network (SAPRIN) has linked established rural nodes with new urban nodes. The goal is to have 1% of the entire population tracked for vital statistics, migration, and many

other issues. The HDSS approach is to annually measure everyone in each space, commonly with a ceiling of 100 000, within official boundaries. The resultant dataset, made freely available by SAPRIN, provides a remarkable picture of each node and their linkages. It also reflects ongoing monitoring and evaluation of service delivery in general, with a strong focus on health as crucial to the transformation of the lives of the impoverished.

The chapter argues that the HDSS approach may circumvent some of the power dynamics that flow from a funder seeking to influence how a monitoring system works, whom they talk to, what questions they ask, and so on, but HDSS is by no means immune to the impact of power. The HDSS strategy shares the same challenges, specifically in urban areas, as any monitoring system: access, trust, and competition for time and attention. The HDSS is another form of data collection, not a rival to other monitoring, and may offer some valuable insights.

Monitoring and HDSSs

Monitoring systems comprising the regular collection and analysis of data have appeared in virtually every aspect of public and private life. According to the UN Task Force (1984:7), both monitoring and evaluation (M&E) have "emerged both at country level and in the UN system in the early 1950s. Since then, it has evolved slowly and unevenly". While this held true in the early 1980s, M&E was ubiquitous by the turn of the century.

As this book clarifies, 'M' and 'E' are available for integration but are also very different activities. Monitoring systems are not always or automatically linked to periodic reviews or evaluations and vice versa; each has its specific role and function and methodological demands, despite being repeatedly conjoined as 'M&E'. Both can exist without the other. Nevertheless, as 'M&E', they have become a global industry, whether named as M&E or performance measurement, or translated into balanced scorecards and results-based reporting schemes, and any number of tools purported to measure and improve human progress.

M&E-related systems, departmental units, and flashy dashboards are now a near-universal requirement in government, much of the private sector, and most of the development sector. Whether this increased spread can be equated with increased efficiency, effectiveness, or impact remains debatable.

Monitoring and the associated language of "M&E" has moved from an essentially pro-poor development location and orientation to being a "management tool". Influential donors and governments rely on "M&E" to provide evidence of a return on their "investments".

Monitoring has become a critical element of private and public sector management (performance measurement). The language of M&E, from 'smart indicators' to 'theory of change', has percolated policy and management circles and is rapidly losing meaning through repeated misuse. 'Monitoring' can also be found at the forefront of the monetisation of lifestyles. Your smartwatch or cell phone measures your performance throughout the day – steps, oxygen levels, heart rate, stress, and mindfulness.

Work-based performance measurement systems will seek to assess your efficiency at the office, using indicators you may or may not have any input into or knowledge of. All of this will be digitally captured by more or less visible cameras that will film it, and digital surveillance - in public spaces – has become literal and ubiquitous. In addition, basic machine learning allows for pattern recognition even in these massive, data-heavy systems, with obvious potential for 'monitoring as policing' to become commonplace. The Chinese 'social points' nod very heavily in this direction (Kobie 2019).

Most M&E practitioners would be appalled to be located in this space but should be aware of what type of monitoring is occurring and for what purpose. This awareness should acknowledge how developmental tools can be used for different purposes. Most of the chapters in this book engage with monitoring in a developmental context, as this chapter does, but the broader reach and consequences of monitoring should be considered.

After primarily being a development sector activity, technology has allowed monitoring to explode into our lives, but the quality of life for the impoverished has degraded. As monitoring and measurement systems have become more complex and their dashboards more colourful, the world has become increasingly unequal; the poor are faced with the direct effects of climate change, pandemics (including COVID-19) and the like. As Evans noted, "In the early years of the 21st century, it is indeed worrisome that in many parts of the world we remain 'in the dark' about life's vital events – namely birth and death" (InDepth Network 2005:7). Monitoring is itself beset by late-capitalist inequality: the rich can measure every breath they take, while the poor may live and die without being officially recorded at all. The politics of monitoring cannot be avoided.

This inequality – more pointedly, the search for health equity that it triggers – lies at the heart of health and demographic surveillance systems (HDSS), which provide accurate, longitudinal coverage of the vital events, demography, and health status of an entire population in each space. As Kahn, Collinson, Hargreaves, Clark and Tollman (2005:88) noted,

> "Health inequities … result from differences in health status outcomes between groups that are avoidable and unnecessary, and hence unacceptable and unjust. Widening health gaps … call for careful monitoring and evaluation of interventions, not just for their impact on health outcomes but also their impact with respect to these 'equity gaps'."

To cite Evans again, HDSS starts from a progressive point of view, that "all lives count and deserve to be counted" (InDepth Network 2005:7). So too, the common assumption of HDSS is that all data should (after suitable anonymisation) be made freely available to anyone seeking to use it for research. HDSSs have yet to fall prey to the marketisation that has befallen much of M&E, not least because HDSSs, by design, are in areas of substantial poverty. Nonetheless, it would be naïve to imagine that any system generating data is not at risk of commercialisation or worse.

M&E practitioners are frequently challenged by their desire to involve community members in participatory monitoring to get local buy-in and better understand the context and local nuance. M&E systems need to account for the rhythms and requirements of local communities and operate at their pace, in the local dialect, with an appropriate understanding of, among other things, local histories, local hierarchies, politics, customs, and seasons. These are essential requirements for successful development. However, they take time and money. Those paying for the programme often want to see 'bricks and mortar' (things being built), not consultations with communities and other activities they do not regard as relevant. This has been described as "a movement away from participatory approaches and towards more top-down control amongst the donor community, with increased emphasis on results-based management" (INTRAC 2020).

There is a broader tension between northern donors and funding agencies and the realities of working in the Global South, and many similar issues affect HDSSs, though in different ways. For example: who says what questions may or may not be asked? Who has the final say over indicator selection? Who performs the analysis? Are recommendations workshopped with the local community or taken over by the donor? Does management take informed decisions that include the views of the community? Who has the power to make or break the system? Valadez and Bamberger's (1994:7) magisterial text on M&E in developing countries included the following, "… the OECD concludes that, with the exception of a small number of countries, 'interest in evaluation generally tends to be stronger among those allocating resources than among those using them'".

The situation may have been different if M&E were a partnership, not an imposition on the developing world by the North. The same is true at a community level; regardless of who is paying for the monitoring, it cannot be imposed on communities if it is intended to be developmental. Power is always at work, in M&E as in HDSSs. Power

matters across design, implementation, analysis, reporting, and dissemination. When power is examined, the dynamics within any M&E system or HDSS become clearer. We may better understand relations between implementers and donors and the multiple actors in any developmental intervention. This is without bringing into the argument issues of indigenous research methods and the need to ensure that the community sets the pace of and is involved in the content of a monitoring system and evaluations (Chilisa 2012). Valadez and Bamberger had noted (1994:404), although without the community featuring:

> "Monitoring is a source of power to those who control the system and a threat to agencies that do not have this control. The potentially threatening nature of evaluation has had a destabilising effect on M/E systems, since those agencies being evaluated have sought to limit the use of any data that affect their budgets and programmes. Furthermore, central agencies may compete among themselves to control the systems."

Among many practitioners, the impulse behind HDSS and M&E is a similarly developmental, progressive, pro-poor approach. M&E is frequently used in impoverished settings, just as HDSSs are deliberately located in places that are poor and vital events are at risk of not being recorded. Both approaches seek to utilise participatory approaches where possible, combined with rigorous data gathering, cleaning, and analysis, to best understand what is working, what is not, why, and how the situation may be improved. In both, where developmental approaches are used, HDSS and M&E practitioners know that communities have their own pace, rhythm, and ways of operating, which are not always amenable to researchers. Both HDSS and M&E practitioners prefer open data, sharing of resources, and collaborative learning. Both need funds to do their work, which is one reason to always look at power dynamics.

HDSS staff are also paid – by the government, in the case being presented here – and so, the power dynamics implicit in any arrangement when one party funds and sets the parameters and another carries out the work also play out in an HDSS. However, they do so differently because an HDSS provides data that is absent or uneven in official statistics (vital statistics, most obviously). In South Africa, the SAPRIN network works closely with Statistics South Africa (StatsSA), providing an important additional set of analytic skills, tools, and training opportunities for the analysis of longitudinal data within StatsSA defined geography, as well as working collaboratively to verify data. Moreover, HDSS funding is of little value if it is not long-term, relying on sound relationships. The M&E field is more likely to be shorter-term contracts for specific programmes or projects, possibly lasting from a few weeks to a few years, but with a predetermined end date. This difference is critical: HDSSs are put in place for medium to long-term monitoring and analysis, and relationships reflect this. Expectations

of HDSSs differ from monitoring systems linked to specific projects, which must account (for example) for monies spent on the project, how many days of work were created, how many people received training, etc. Monitoring is at the heart of both but plays out in different ways.

There can be a different tone to the relationship, in that (in this case) a domestic government is funding the HDSS to advise them as to what interventions (primarily health-related) are working and which are not, over time, as reflected by the population under study. SAPRIN is continually assessing and feeding data and analysis back into government systems, but unlike a programme-linked system, the HDSS is more a partner than a contractor. This is a deliberate intervention to improve the lives of the poor by developing detailed household profiles as well as linking these to official records such as child school attendance and road to health cards. For example, the HDSSs utilise Verbal Autopsies, which have proved vital for capturing things "likely to be particularly sensitive to recall bias" (Sankoh and Byass 2012:580). On precisely this issue, the Dande HDSS in Angola found that:

> "The neonatal, infant, and under-five mortalities reported by the National Institute of Statistics of Angola for the years 2011 – 2015 are 24, 44, and 68 per 1000 for the country, respectively. The total fertility rate for the period 2013 – 2016 is 6.2. Compared with the Dande HDSS data, a probable under-reporting of neonatal deaths is perceived. These events are very likely not to be reported some time after their occurrence, particularly if the child was not listed in the household registration book. Also, cultural aspects or embarrassment of respondents to talk about deceased relatives might contribute to the omission of death events. This is in line with previously published data and has been documented as one of the reasons why child mortality is most probably underestimated in many surveys in sub-Saharan Africa. Implementation of new projects around maternal and child health are planned, including a birth cohort which will enable accuracy and a thorough knowledge of these events" (Rosário, Costa, Francisco and Brito 2017:4).

Tension between programme partners is not always a bad thing and can lead to debate, discussion, and a better way forward. In an HDSS, relations between funder and implementing agent are quite specific. The HDSS is uniquely positioned to work with the government as a long-term partner, filling data gaps, assessing policy impact, adding nuance, and providing unparalleled insight into its population.

HDSSs are not cheap. Fieldwork continues for 48 weeks of the year, and very powerful databases are required to run multi-year analyses. HDSSs also design and launch smaller studies within the node, as the quotation suggests, to try and better understand the dynamics around an issue, in this case, as painful and culturally

sensitive as neonatal mortality. An HDSS can thus commission an evaluation if it wishes, which indicates a significant difference: HDSSs are institutional and have the autonomy to carry out their own evaluations or add to their monitoring questions. Smaller studies can be on-boarded in an HDSS because all HDSSs start with a census of their node. In return, year on year, to the same dwelling and (most of) the same respondents. This kind of local, detailed, longitudinal knowledge – encoded in data rather than locked into the minds of local leaders – makes HDSS sites attractive for smaller, sampled, and focused studies.

Case study methodology talks to the power of the case because the depth of understanding of the case (in this instance, the node) is unparalleled. However, because the case is about depth, not breadth, generalisability is unavailable (Yin 2018). The case may include local exceptions that will not be evident to the researcher due to them exclusively working in that case site. However, an HDSS can be thought of as a longitudinal, quantitative (and rather large) case study. It may be more helpful than comparing it with sample surveys or an integrated M&E system. A key issue facing many HDSS proponents is the lack of national representivity in the data (Byass, Sankoh, Tollman, Hogberg and Wall 2011). The findings cannot be generalised to a larger population, something it shares with case study methodology. It may seem odd, given that a sample of a thousand respondents can be used to generalise, within measurably accurate parameters, across the entire country, but an HDSS with a population of 100 000 under surveillance cannot. That, of course, is the benefit of sampling. Nevertheless, the HDSS has the power of an ongoing case study, resting on uniquely deep, nuanced, and longitudinal data for an entire population in each area, falling within official boundaries.

Moreover, in South Africa, SAPRIN has a network of HDSSs, surveilling hundreds of thousands of people every year, year after year. The data is merged into a single dataset, and the complaint that "conventional statistical theory is not particularly useful for addressing this question" is true of any stand-alone HDSS (Byass, Sankoh, Tollman, Hogberg and Wall 2011:1). However, improvement of statistical language may be needed to understand the power and value of this network of longitudinal, large-scale cases, where all respondents are being asked the same questions in the same way, every year. The power of this dataset does not lie in national representivity but depth and complexity, and above all, accuracy.

HDSSs generate a geospatial coordinate map of every bounded structure, then every dwelling, number of households per dwelling, vital statistics per household, as well as socioeconomic status and a raft of other questions. A Hillbrow apartment block could be considered a bounded structure; each flat is a dwelling. Within that dwelling may be a single person, groups of individuals, a household, or multiple households. All this needs to be recorded and updated annually. The focus of an HDSS is to improve the lives of the poor, and all the questions are geared at understanding health and

socioeconomic status and assessing health and population changes. These are asked year on year so that data errors can be identified (or seen as emerging social phenomena) and corrected. So, an HDSS does not claim survey-style representivity – if that means 'the entire nation' – but the SAPRIN nodes analysing data on hundreds of thousands of poor communities, households, and individuals does mean that SAPRIN can make the case that the network is indeed representative of the poor, and as a living dataset that can self-correct via annual implementation. Aside from the sheer weight of data, the rigour of the methodology makes SAPRIN a key player in the South African data ecosystem.
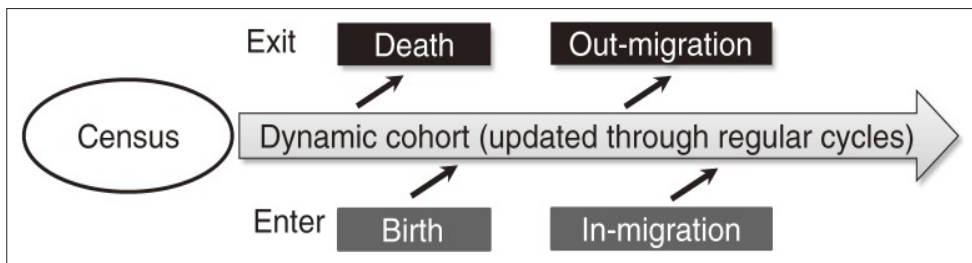

Figure 8.1:   HDSS design (generic)

In a network of interlinked HDSS nodes, such as SAPRIN, the combined power of the same instrument being applied to vast numbers of people in different contexts and geographical locations suggests a unique data power, particularly when several nodes are linked, and their data is merged. This is pertinent to the added possibility of the return to precisely the same respondent within months, and again face-to-face within a year with follow-up diagnostic questions and trials.

HDSS, Monitoring and Incentives

An HDSS has the core function of measuring a given population using multiple interactions each year so that respondents keep them apprised of the entire course of significant events, such as pregnancy. This, in turn, demands that an HDSS is underpinned by robust community engagement and mobilisation, without which respondent fatigue would be rapid and inevitable. Many clinical trials offer respondents an incentive, and respondents may come to expect that this is a research norm. Mixing non-remunerated fieldwork with smaller on-boarded trials, which may offer incentives, can cause confusion and hostility, as respondents do not understand the difference between the two.

This must be handled with care. The HDSS is, by design, located within a poor node. Resources matter, particularly if respondents feel they are being denied something

they see or hear others receiving or enjoying. Unlike survey field workers, those from the HDSS work permanently within the node. Field workers have project uniforms, tablets for capturing data, vehicles that transport them; all aspects of the HDSS, as of any development intervention, suggests that it is resource-rich, which it is within the local context. For example, in the case of GRT-INSPIRED, Hillbrow – with the highest local population densities in Africa – is a favourite site for research from all over the world, with its mixture of density, bustle, edginess, cosmopolitan culture, and overcrowding. This means that HDSS staff encounter respondents who got paid to participate in a clinical trial and who then question the lack of payment from GRT-INSPIRED. Balancing freedom to research with an appreciation of others working in the same space becomes an issue in urban areas, but less so for rural nodes.

This challenge of well-funded research activities in a poor community goes beyond the issue of clinical trial incentives. In the Nairobi HDSS, for example, the number of formal and informal gatekeepers demanding 'goodwill' to allow fieldwork to continue unabated simply grew too large, once some 'goodwill' had been dispensed (Kyobutungi 2020). The only way a successful HDSS can operate and remain operational is through continued community mobilisation, so that community leaders and members (commonly formed into a Community Advisory Board or CAB) are not merely informed about the project but are key to the feedback loop. The CAB is also able to help identify local contestation, assist with access (which is key for an urban node), and act as ongoing advisors regarding the community. Given the mobility of urban populations, it is important to keep revisiting communities and reminding them of what the HDSS is trying to accomplish for them.

HDSSs grew mainly from the Community-Oriented Primary Health Care (COPC) movement. The COPC approach (like the HDSS) works in defined, poor areas and partners with communities and forms CABs. The orientation is proactive prevention, with the overall goal being support treatment and care, the prevention of disease, and the identification of people with health needs, as well as the assurance that those needs are met. The focus is on the local; in the case of our University of Pretoria colleagues, who run part of the HDSS described here, Community Health Workers accompany field workers during field work, under the guidance of a nurse, so that the HDSS is both 'giving' and 'taking'. Urban respondents are leery of being interviewed at all, let alone three times a year, and the creation of a virtuous cycle that reaps tangible benefits for respondents in return for data is essential. A COPC orientation seeks to make the relationship more balanced. No incentives are paid, but health assessments in particular take place before interviews, and, where possible, results are immediately provided to the respondent.

Moreover, the HDSS nodes in SAPRIN seek to be proactive. It may be argued that an HDSS must be proactive to remain in place with community support. The goal of data collection is not simply the production of academic papers or datasets. Even

without formally adopting COPC, the HDSS seeks to link to official records, such as clinic records, child inoculation, and school attendance, and to provide an evaluation of the effects, or lack thereof, of government interventions to aid the poor. The data should allow local and provincial governments to target specific, identified challenges within the node. The node can aid in the identification of children who do not attend school, are not inoculated, etc., and can assist the relevant adult regarding the necessary steps to be taken. Most HDSSs also collect, analyse, freeze, and store blood samples (usually dry blood spots), meaning that they can be re-analysed years later if required. The HDSS is thus a key meeting point for the community, community leaders, academics, policymakers, and programme delivery managers.
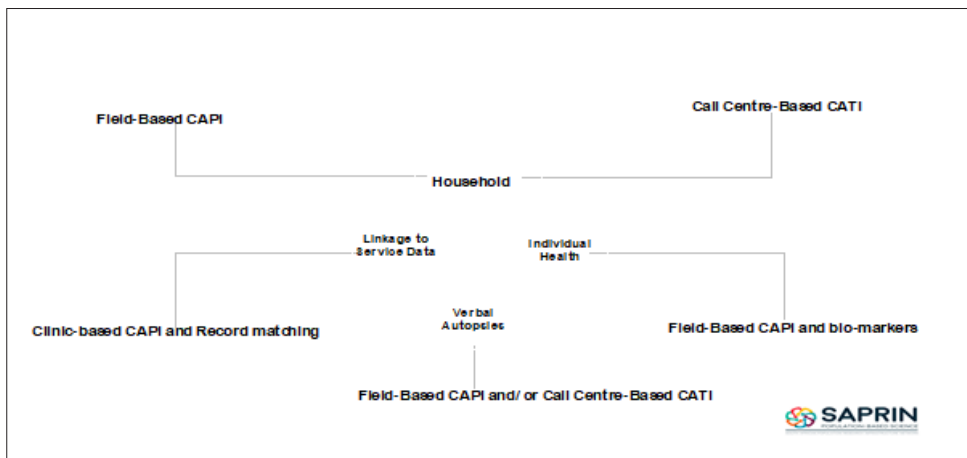

Figure 8.2: SAPRIN, *The SAPRIN model*

As Figure 2 clarifies, the considerable challenges of fieldwork rise to meet another challenge, linking the records of each person in the database with official records such as social security, health records, road to health cards for children, and school attendance.

Verbal autopsies (VA) are unavoidable in the context of poverty and poor record-keeping – although accurate, cause of death data is not merely a challenge for low and middle-income countries (LMICs). As Erin, Nichols, Byass, Chandramohan, Clark, Samuel, Flaxman, Jacob, Leitao, Maire, Rao, Riley, and Seftel (2018:1) noted, on behalf of the WHO Verbal Autopsy Working Group,

> "In low-income countries, many deaths are unregistered, unrecorded, and unnoticed by the health system. Nearly half of all countries fail to meet United Nations standards for death registration (90% coverage), while high-quality cause-of-death data are lacking for 65% of the world's population. Inadequate data on cause-specific mortality

patterns impede the development of sound health policy, planning, monitoring, and evaluation."

Verbal autopsies – based on questions about symptoms shown by the respondent – are now automated, with algorithm-driven diagnoses available. VAs are the best that is available, as Nichols et al. concluded:

> "Despite all improvements in design and technology, VA is only recommended where medical certification of cause of death is not possible. The method can nevertheless provide sufficient information to guide public health priorities in communities in which physician certification of deaths is largely unavailable."

HDSSs seek to influence policymakers by showing policy impacts over time, enhancing service delivery in the HDSS node by identifying health and other needs, and linking respondents with appropriate local service providers, clinics, or social security offices. Using nodes as locations for clinical trials directly impacts the node and society more generally.

HDSSs in South Africa

According to the literature, the first HDSS was established in South Africa in April 1940 (Tollman 1994) at Pholela Health Centre. However, historian Shula Marks has noted earlier Chinese (and other) experience at work; the founders of Pholela had been working in China in the 1930s, which influenced the South African experience (Marks 2013). Using a COPC approach, the HDSS was established by the Ministry of Health to identify and prevent diseases common in rural (then) Natal, such as tuberculosis, smallpox, typhoid, and measles, and to impact health policy (Ye et al. 2012:11). As Marks notes, the use of the term 'social medicine' and 'socialised medicine' in the inter-war years was based on an approach that emphasised the social determinants of health and disease, was inherently political, and influenced by socialism. According to Kark and Abramson (2003:882),

> "South Africa in the 1940s had become a major conceptual leader in the development of what later became known as community-oriented primary care, or COPC ... as an effective mode of delivery of health services to the population. This found expression in the establishment in 1940 by the Health Ministry, ably led by Eustace Cluver and Harry Gear, of a demonstration health centre in a rural Zulu community (the Pholela Health Centre which Sidney Kark headed), the Gluckman Commission's far-seeing recommendation in 1945 of a nationwide network of community health centres..."

The Gluckman Commission had made some powerful recommendations, most obviously for a national health service (influencing the later Beveridge Report in Britain) and an HDSS network. However, support waned once the political tide began to change in the immediate post-war period. When the Nationalist Party won power in 1948, that was the death knell for the Commission's work, but the recommendations were cited again in the 1990-1994 negotiations phase, suggesting that only a democratic state might finally be able to deliver on Gluckman, half a century later (Harrison 1993).

HDSSs were initially stand-alone, but in 1998 the International Network for the Demographic Evaluation of Populations and their Health in Developing Countries (InDepth) network was formed, with 36 member centres running 44 HDSSs. Of those, 32 were located in sub-Saharan Africa (Ye, Yazoume, Wamukoya, Ezeh, Emina and Sankoh 2012:12). By 2017, InDepth had 47 HDSS sites, following roughly three million people. One function of InDepth was to gather data across HDSS and establish standards for data acquisition (Herbst, Juvekar, Jasseh, Berhane, Nguyen Thi Kim Chuc, Seeley, Osman, Clark and Collinson 2022). The primary purpose of the HDSS remained observation of population dynamics in a specific geographic area to support epidemiological and other interventions. This gives some sense of the scales at work: from stand-alone HDSSs to a network including three million people in poor areas.

All data collection has weaknesses. The challenge is to identify and control these weaknesses so that survey data can offer inferential data (the search for generalisability) within measurable degrees of accuracy. Thus, while a census will give the best possible coverage, they occur at best once a decade. Sample surveys occur in-between, in multiple areas, and panel studies based on the core idea of returning to the same respondents are few and, by contrast with SAPRIN, limited in scale (the National Income Dynamics Study (NIDS) is a notable exception). Surveys are inevitably a trade-off between depth and breadth, usually governed by cost, and their results are a trade-off between accuracy and generalisability.

NIDS (using pre-COVID-19 figures) interviewed 39 400 individuals, in 10 800 households, in 2017 – the fifth outing for the study allowing face-to-face field work. SAPRIN annually gathers data on hundreds of thousands of people living in poor areas. SAPRIN uses GIS to code and map every structure in the node so that each node has a live map of every structure and specifies the multiple functions of a given structure. In urban areas, a dwelling may also be a car repair site but also a tuck shop, and may offer hair braiding, and cell phone repair. This all must be captured and combined with questions about transport, economic opportunities, etc. Once GRT-INSPIRED is fully operational (in 2022/3) and the Cape Town node (C-SHARP) a year later, SAPRIN will generate data annually on about half a million people living in poor rural and urban areas.

Urban Fieldwork Challenges

Migration coupled with the legacy of the past has demographic effects that make Gauteng appear different than anywhere else in South Africa – it has more men than women, and almost 30% of all youth in the country (5.10 million or 28.6%) live in Gauteng (StatsSA, 2021). Far fewer older people live in Gauteng than in other provinces. This suggests that the health and demographic profile of Gauteng will differ from other, less urbanised provinces. The challenge is building trust to gain long-term access, which is made challenging by the mobility of many residents. There are many more attractions in urban areas than answering questions or giving blood samples.

In the era of big data, when 'harvesting' unverified secondary data is seen as cost-efficient, and thus attractive, the HDSS approach reminds us of the importance of accuracy in data, not the flashy dashboard, and of the value of face-to-face field work based on good community relations. This is true for all monitoring systems based on face-to-face interactions with people, which generate primary data that are measurably accurate. Repeat interventions in the same space with the same respondents build relations and improve data quality. A respondent may only open up about some vital event after a few years when they have come to trust the annual interview more or have seen benefits from participating. They may point out where they have been misleading field workers in prior interviews out of embarrassment or choice. An HDSS can make retrospective changes, which is part of the 'monitoring of a special type' an HDSS enjoys.

There is massive competition for peoples' time and attention in urban areas. In a space such as Hillbrow, there are multiple surveys, monitoring, clinical trials, and other interventions taking place. The environment combines a street economy of hustling and a 'flatland' that mixes captured buildings that are unsafe to enter, others where field workers are simply refused entry, and others that welcome field work. Access is always a challenge in urban areas, but it has led to new partnerships and may, in time, generate the need for smartphone applications to gather and send data. The sheer demographic size of Hillbrow makes it very difficult not to be sampled by any survey agency working in Gauteng. The HDSS must find purchase in this contested space, where field workers face security challenges and where respondents are preoccupied and streetwise. This contrasts with the HDSS in Morogoro, in rural Tanzania, which found that national and district officials had over-estimated numbers attending clinics for malaria treatment. The area had a 30% mortality rate because of malaria, which rose to 45% in children aged below five. Due to the HDSS data, a five-fold increase in malaria control resources was made available, and a 20-fold decrease in the under-five population (Ye, Yazoume, Wamukoya, Ezeh, Emina and Sankoh 2012:14). Any monitoring system operating in a

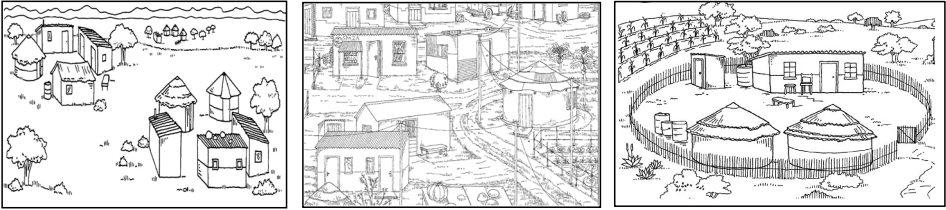cosmopolitan urban area such as Hillbrow has very little chance of such a neat match of need and supply.

Figure 8.3:   SAPRIN, *SAPRIN 'bounded structure' images*

As noted above, an HDSS is created first by completing a census of physical structures and households in the area as a baseline, followed by regular visits to each household to gather health and demographic data. The cohort is dynamic in that members are added through birth or immigration or leave the cohort through death or emigration. Tracking population migration is particularly important for HDSS data (Ginsburg, Collinson, Gómez-Olivé, Gross, Harawa, Lurie, Mukondwa, Pheiffer, Tollman, Wang and White 2021).
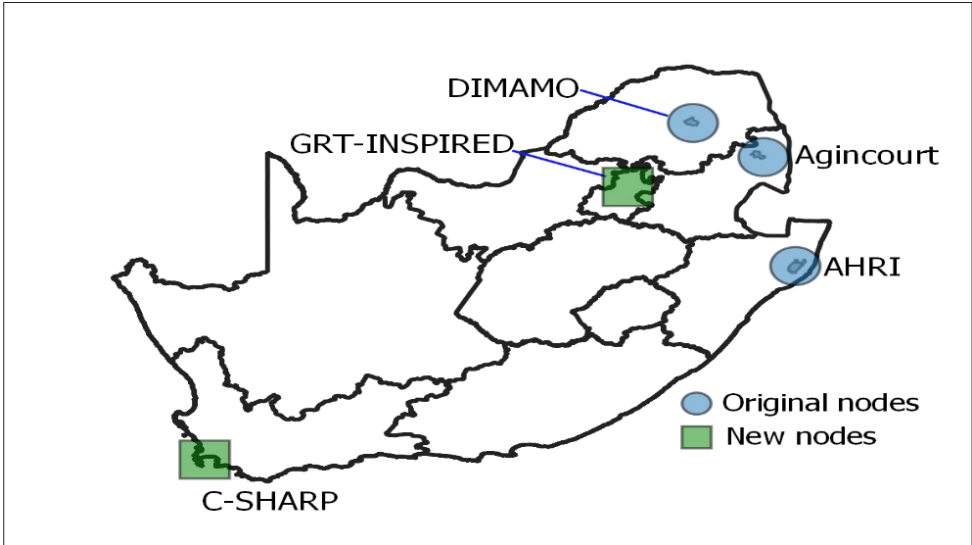

Figure 8.4:   SAPRAN, *SAPRIN nodes*

As Figure 4 suggests, SAPRIN plans additional nodes in the Eastern Cape and eThekwini. As a result, migration should become an even more important work area for SAPRIN.

A Quick Overview of GRT–Inspired

The Gauteng Research Triangle is a partnership between three major universities, each taking responsibility for different sites or parts of sites. The University of Pretoria, acting via the Department of Family Medicine, is responsible for all fieldwork in Melusi and Atteridgeville. Together, they generate data on 50 000 people. The University of Johannesburg and the Wits Reproductive Health Institute have divided the Hillbrow sub-place roughly in half. Each is responsible for data on 25 000 people, making up the other half of the 100 000-strong node. The selected Small Area Layers from StatsSA form a single contiguous space for each site. These three sites comprise the node.

Sankoh and Byass (2012:2) noted the following challenge:

> "In practical terms, one important consideration is whether the final population is defined as being within a contiguous area or in a collection of small areas (e.g., discrete villages) within a wider area. This has important logistic implications in terms of organising and maintaining on-going surveillance, as well as affecting the definition of migration events…"

The exact impact of these three separate sites remains to be seen. In Nairobi, the HDSS was in two slum sites but is currently redesigning to move into more spaces (Kyobutungi 2020). Given the complexity of the urban form in Gauteng, it was deemed important to spread the sites comprising the node to cover different spatial areas within Gauteng, but also very different urban forms.
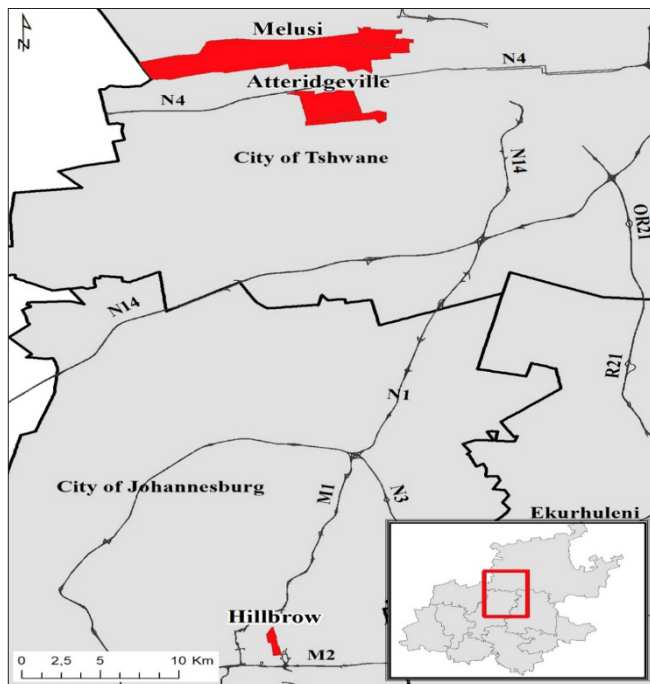
Chapter 8

Figure 8.5:   GRT-INSPIRED, *The three sites of GRT-INSPIRED (sub-places)*

The node is entitled 'GRT–INSPIRED: The Gauteng Research Triangle Initiative for the Study of Population, Infrastructure, and Regional Economic Development' because the approach from design onwards was multidisciplinary. Health and demography underpin the entire strategy but are integrated with urbanism and the built environment to understand the relationships between health outcomes and inequality, class, socioeconomic status, and related issues. Climate change is also key, and the relationship between it and health, demographics, and economics is a cross-cutting concern. GRT-INSPIRED will be installing thermometers in 10% of sampled dwellings to measure temperature change generally as well as extreme weather events and their relationship with disease. The critical issue of migrancy – internal and (in stark contrast to rural nodes) international – will be a significant focus.

In contrast with the rural homestead sketch above, Frith (s.a.), using Census 2011, calculated Hillbrow (the sub-place) was only 1.08km² yet had a population of 74 131 at a staggering density of 68 418/km².

As Figure 6 clarifies, densities are high in much of Gauteng, especially showing apartheid spatial architecture, but all are dwarfed by Hillbrow (the higher the spike

on the image, the greater the density). If Hillbrow stands out due to density and sheer volume, Atteridgeville, the second site, is an old apartheid township. Like most townships, the small gardens of formal structures now mostly boast additional living quarters (for rent), ranging from formally built outside rooms to garages, tents, shacks, caravans, and abandoned old cars. People will do what they must to gain a toehold in the city.
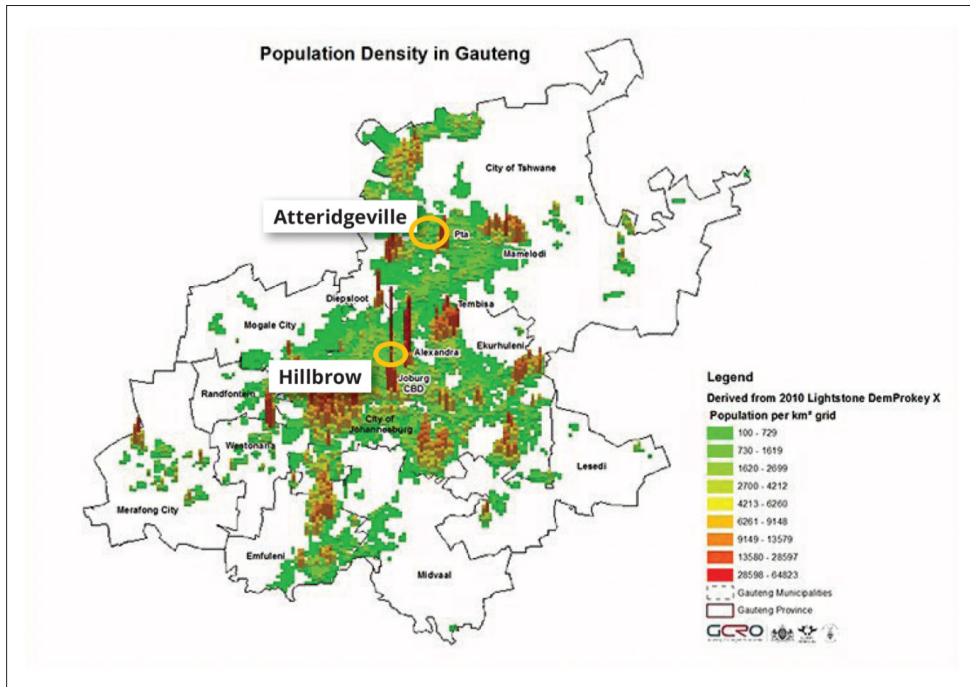
Figure 8.6:   GCRO, *Population densities (and the GRT sites)*

The need for a toehold in the city is made clear in a different way in Melusi, which did not appear in Census 2011. However, by 2021 it had an estimated population of over 40 000 people, and the area is becoming a site and service settlement. The University of Pretoria (UP) had begun working in the area using a shipping container that was re-purposed as a community clinic, and Community Health Workers (CHWs) were deployed in the area by the provincial government to work with UP in providing COPC. Those CHWs are now part of the field team, accompanying field workers, and more proactively link residents with available services.

Figure 8.7: University of Pretoria, *Melusi*

The map, using official boundaries (sub places), does not give a sense of what is happening on the ground in Melusi. The Google Earth image below paints a somewhat different picture:

Figure 8.8: Google Earth, *Melusi with detail*

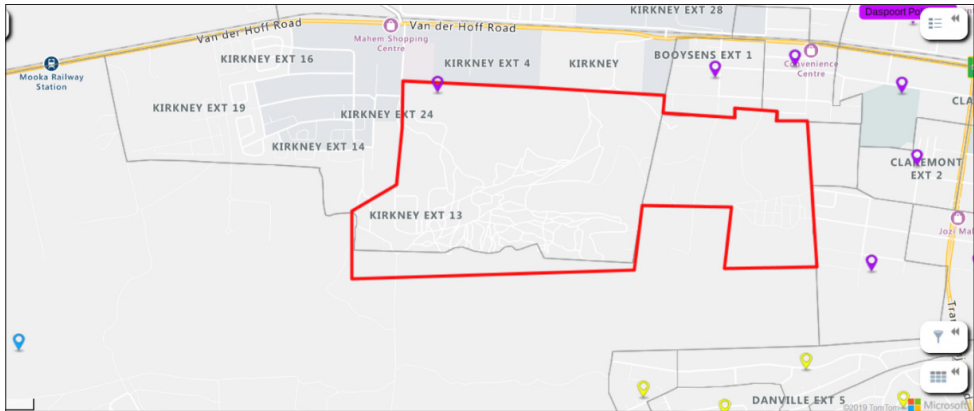The use of unclean water from the nearby abandoned quarry is self-evident from the image, which underscores the importance of the government as a partner in this intervention. If services are not provided, especially reticulation of potable water and sanitation, the pollution and disease that will follow (and are already present) are clear. However, the provision of services often attracts further growth. Gauteng grows at some 2.6% per year in population, as services and opportunities attract people. Policy interventions to better understand how to manage informality are inseparable from policy interventions to improve health outcomes.

Another reason for splitting the node into sites, aside from capturing different urban forms as well as different parts of Gauteng geography, came from the *South African Research Infrastructure Roadmap* (Department of Science and Technology 2016:35), which notes that South Africa faces "several challenges, including high levels of inequality ... an unemployment rate of 25% [now 29%] ... a poverty headcount ratio of 57% as well as colliding epidemics of HIV/TB and non-communicable diseases". In other words, separating any single issue from the 'colliding' challenges concentrated in Gauteng may miss a more complex picture, particularly of cause and effect. Gauteng is the smallest province in South Africa but has the largest population share (at 26% - 15 176 115 people), generating 34% of the national GDP (Katumbe and Everatt 2021).

The poverty headcount ratio is far lower (29,3%) than nationally, but the province has 1.9 million people living with HIV, second only to KwaZulu-Natal. The province includes Johannesburg, the most unequal city on the planet (Economist 2021). As StatsSA (s.a.) noted some years ago, Gauteng's economy is roughly the size of Morocco's national economy. It is the 7th largest in Africa but is far from equally shared.

Demographic surveillance in Gauteng must consider the province's urban, demographic, socioeconomic, and other characteristics. Gauteng is not simply 'urban' – it is a highly complex, profoundly unequal, racialised, and fragmented space comprising three metropolitan areas that form a continuous urban area. Gauteng offers a panoply of urban forms exacerbated by inequality and race, although the HDSS is not required in gated communities, lifestyle estates, or suburbia. A multi- or transdisciplinary approach and teams of analysts that locate HDSS longitudinal data in the varied contexts Gauteng offers, focusing on poorer areas, is vital. GRT-INSPIRED needs to locate health and demographic change in the messy realities of urbanism, which is no small task.

Our partners from the University of Johannesburg (UJ), for example, bring a different kind of science via their Water and Health Research Centre (WASH) and opportunities for non-health students to get into the field. UJ is a field partner for the SAPRIN protocols and has an additional intervention analysing the use, storage, and safety of stored water and sanitation. This involves getting access to high rise apartments in

Hillbrow and the various forms of sanitation in Atteridgeville (formal and backyard), and Melusi. This is also important for community education. In the Angolan Dande HDSS, a water, sanitation, and hygiene (WASH) programme was launched within the node to reduce anaemia and malnutrition, and to protect preschool children (Rosario et al.). In coastal Kenya, 'pop up' questions, in addition to the protocol, include water sources and health (Kaneko, K'opiyo, Kiche, Wanyua, Goto, Tanaka, Changoma, Ndemwa, Komazawa, Karama, Moji and Shimada 2012:5). With COVID-19, this area of work became even more critical. Many households are still obliged to share sanitation facilities. Even a shared cloth for cleaning may become an agent of transmission, as with water stored in open containers. This also provides immediate feedback to the respondent.
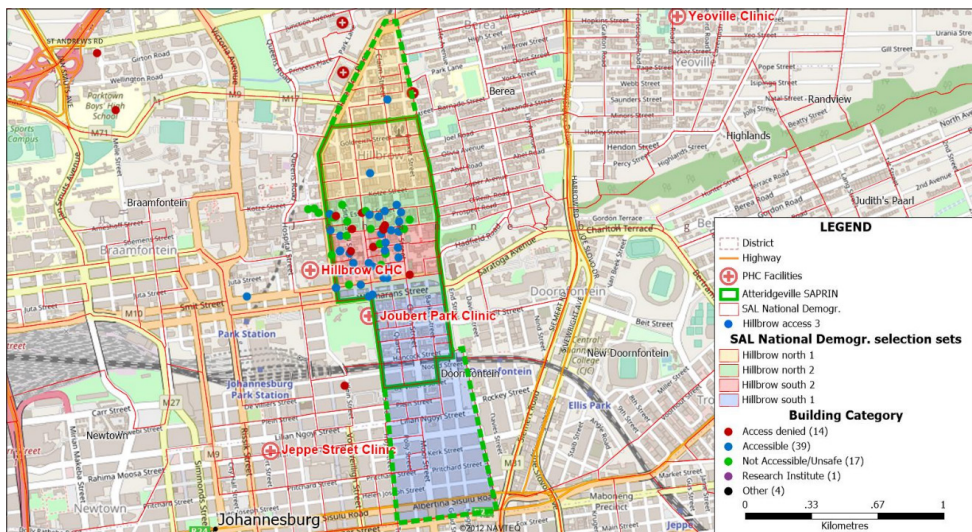
Figure 8.9: University of Pretoria, *The Hillbrow sub-place*

The challenge of access was expected to be primary in an urban HDSS, exactly as it is for survey or monitoring system fieldworkers, for reasons outlined above. Our Wits partner, the WRHI, has been working in Hillbrow since 2002, using CABs and other forms of community engagement. Given Hillbrow's reputation for criminal activities combined with massive diversity and similar population density, WRHI needed both a physical and reputational presence.

To contrast with the uniqueness of Hillbrow's high rise buildings and the informality of Melusi, Atteridgeville was selected as a classic 'township' designed by the apartheid regime to house black Africans who had managed to win rights to remain in the urban space. After democracy, many township households split, because, for the first time, there were no race-based restrictions on where they could live. Younger

members frequently took the first step out of home into an informal settlement (Everatt, Jennings and Gotz 2005). The density challenge was somewhat reduced by the ability to move, even though economics firmly dictated where people could (afford to) move.

In Atteridgeville, GRT-INSPIRED found formal dwellings with various other structures in their backyards. Every household in every structure must be enumerated. Access was less of a challenge because the UP Family Medicine Department has been working from Kalafong Hospital in the area for years. Community mobilisation is a critical mainstay of GRT-INSPIRED in all areas but has been most needed in Atteridgeville. The layout makes planning somewhat more manageable, as Figure 10 suggests.
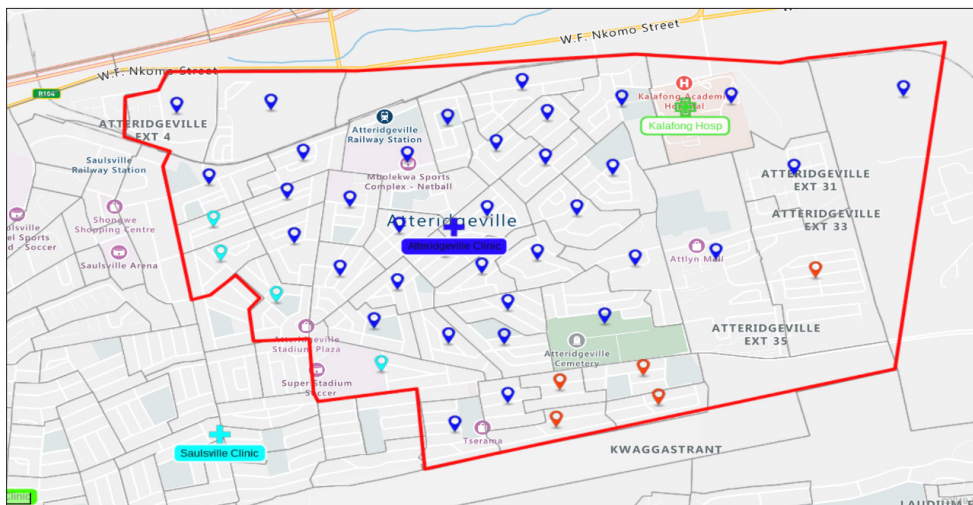
Figure 8.10: University of Pretoria, *Mapping fieldwork in Atteridgeville*

Conclusion

Monitoring is not value neutral. Monitoring is done with a purpose, which generally involves reaching evidence-based conclusions that can inform programmes, policy, and planning. The chapter opened by warning of the dangers, in a monitoring system or an HDSS, of being misused (such as digital imagery) or hi-jacked to become a commodity. Monitoring, as discussed here, is developmental. It is in place to help impoverished communities by providing some immediate benefits (such as COPC), but also for the long-term health of the community by linking findings to official records and partnering with the government to identify key areas of intervention that may be needed. The HDSS will still be in place after that intervention and can then evaluate its efficacy for the next cycle.

HDSSs are not a rival or a substitute for monitoring or M&E systems. It is a different approach to data and understanding population dynamics, as well as development priorities, rather than being linked to a specific programme or project. However, all have the same goal, which is to work for improved quality of life in poor areas. This requires the government to pay attention to the findings, and strong governance to ensure that benefits intended for the poor reach them to help move people out of ill health and poverty, and towards the South Africa envisioned in 1994.

Reference List

Byass, P. Sankoh, O. Tollman, S.M. Hogberg, U. & Wall, S. 2011. Lessons from History for Designing and Validating Epidemiological Surveillance in Uncounted Populations. *PLoS ONE* 6(8): e22897. https://doi.org/10.1371/journal.pone.0022897

Chilisa, B. 2012. *Indigenous Research Methodologies*. Sage, London.

Collinson, M., Mudzana, T., Mutevedzi, T., Kahn K., Maimela, E., Gómez-Olivé F., Gomez, Mngomezulu, T., Gareta, D., Kabudula, C. W., Nemuramba R., Tlouyamma, J., Tollman, S. & Herbst, K. 2021. 'Cohort Profile: South African Population Research Infrastructure Network (SAPRIN)' in *International Journal of Epidemiology*. https://doi.org/10.1093/ije/dyab261

Department of Science and Technology. 2016. *South African Research Infrastructure Roadmap.* Pretoria: Department of Science and Technology.

Economist. 2021. Unpicking inequality in South Africa. *The Economist Middle East and Africa*. https://www.economist.com/middle-east-and-africa/2021/09/23/unpicking-inequality-in-south-africa.

Everatt, D., Gotz, G. & Jennings, J. 2005. *Living for the sake of living': Partnerships between the Poor and Local Government in Johannesburg.* UNRISD Democracy, Governance & Human Rights paper 16, Geneva.

Frith, A. *Census 2011.* https://census2011.adrianfrith.com/

Ginsburg, C., Collinson, MA., Gómez-Olivé, FX., Gross, M., Harawa, S., Lurie, MN., Mukondwa, K., Pheiffer, CF., Tollman, S., Wang, R., & White, MJ. 2021. Internal migration and health in South Africa: determinants of healthcare utilisation in a young adult cohort. *BMC Public Health.* 2021 Mar 20;21(1), p. 554. https://doi.org/10.1186/s12889-021-10590-6

Harrison, D. 1993. The National Health Services Commission, 1942 - 1944 - its origins and outcome. *SAMJ* 1993; 83, pp. 679-684.

Herbst, K., Juvekar, S., Jasseh, M., Berhane, Y., Nguyen Thi Kim Chuc, Seeley, J., Osman, S., Clark, S. J. & Collinson, M. 2022. Health and demographic surveillance systems, history, state of the art and future prospects. Special Issue of *Global Health Action* in Honour of Peter Byass. https://doi.org/10.1080/16549716.2021.1974676

InDepth Network. 2005. Measuring Health Equity in Small Areas: Findings from Demographic Surveillance Systems. Routledge & CRC Press. https://www.routledge.com/Measuring-Health-Equity-in-Small-Areas-Findings-from-Demographic-Surveillance/Network/p/book/9780367667269

INTRAC. 2020. 'Participatory M&E' https://www.intrac.org/wpcms/wp-content/uploads/2017/01/Participatory-ME.pdf

Kahn, K., Collinson, M., Hargreaves, J., Clark S. & Tollman S. 2005. Socioeconomic status and child mortality in a rural sub-district of South Africa. *InDepth Network: Measuring Health Equity in Small Areas: Findings from Demographic Surveillance Systems.* Routledge & CRC Press. https://www.routledge.com/Measuring-Health-Equity-in-Small-Areas-Findings-from-Demographic-Surveillance/Network/p/book/9780367667269

Kaneko, S., K'opiyo, J., Kiche, I., Wanyua, S., Goto, K., Tanaka, J., Changoma, M., Ndemwa, M., Komazawa, O., Karama, M., Moji, K. & Shimada, M. 2012. Health and Demographic Surveillance System in the Western and coastal areas of Kenya: an infrastructure for epidemiologic studies in Africa. *J Epidemiol.* 2012;22(3), pp. 276-85. https://doi.org/10.2188/jea.JE20110078

Kark, J.D. & Abramson, J. H. 2003. Sidney Kark's contributions to epidemiology and community medicine. *International Journal of Epidemiology*, Volume 32, Issue 5, October 2003, pp. 882–884. https://doi.org/10.1093/ije/dyg291

Kobie, N. 2019. The complicated truth about China's social credit system. *WIRED*. https://www.wired.co.uk/article/china-social-credit-system-explained

Le Roux, H. 2014. The Ethiopian Quarter. *Changing Space, Changing City: Johannesburg after Apartheid - Open Access Selection*. edited by Philip Harrison et al., Wits University Press, 2014, pp. 498–505. https://doi.org/10.18772/22014107656.31

Marks, S. 2013. Social Justice or Grandiose Scheme? Seminar paper presented at WISER, September 2013. https://wiser.wits.ac.za/system/files/seminar/Marks2013.pdf

Nichols, E.K., Byass, P., Chandramohan, D., Clark, S.J., Flaxman, A.D. & Jakob, R., et al. 2018. The WHO 2016 verbal autopsy instrument: An international standard suitable for automated analysis. InterVA, InSilicoVA, and Tariff 2.0.' *PLoS Med* 15(1): e1002486. https://doi.org/10.1371/journal.pmed.1002486

Rosário, E., Costa, D., Francisco, D. & Brito, M. 2017. HDSS Profile: The Dande Health and Demographic Surveillance System (Dande HDSS, Angola). *International journal of epidemiology*, *46*(4), pp. 1094–1094g. https://doi.org/10.1093/ije/dyx072

Sankoh, O. & Byass, P. 2012. The INDEPTH Network: filling vital gaps in global epidemiology. *International Journal of Epidemiology*, Volume 41, Issue 3, pp. 579- 588. https://doi.org/10.1093/ije/dys081

Tollman, S.M. 1994. The Pholela Health Centre - the origins of community-oriented primary health care (COPC). An appreciation of the work of Sidney and Emily Kark'. *SAMJ* 1994 Oct; 84(10):653-8

United Nations Task Force. 1984. *Guiding principles for the design and use of monitoring and evaluation in rural development projects and programmes*. https://digitallibrary.un.org/record/1079?ln=en

Valadez, J. J. & Bamberger, M. 1994. *Monitoring and evaluating social programs in developing countries: A handbook for policymakers, managers, and researchers*. Washington, D.C.: World Bank. https://doi.org/10.1596/0-8213-2989-8

Ye, Y., Wamukoya, M., Ezeh, A., Emina, J.B.O. & Sankoh, O. 2012. Health and demographic surveillance systems: a step towards full civil registration and vital statistics systems in sub-Saharan Africa? *BMC Public Health* 12, p. 741, 2012. https://doi.org/10.1186/1471-2458-12-741

Yin, R. K. 2018. *Case study research and applications* (6th ed.). Thousand Oaks: SAGE Publications.

Chapter 9: Public Sector Monitoring Systems in Practice: The Case of Benin

Abdoulaye Gounou and Damasse Soussou

Introduction to Performance Monitoring Systems

Despite many changes because of reforms over 15 years, the core mandates of Benin's monitoring systems have remained relatively constant: to promote the institutionalisation of evidence use across all levels of government for improved service delivery and development. Over the years, this core function was located at the highest level of the centre of government, to ensure political championing and uptake of evidence use across ministries and departments in Benin. The increasing demand for evidence production and use, however, still requires significant work in terms of capacity development and putting in place systems and institutions that can work across the whole of government. For example, terms and concepts related to results are not consistently applied across the government system. Additionally, despite concentrated efforts by senior government officials to build a system that meets the knowledge needs of different stakeholders, monitoring is still being used primarily for compliance and control. In some parts of the government, it has been possible to foster a culture of learning, but there are still areas where monitoring generates perverse incentives to game the system for the purposes of compliance. A current priority is to systematise capacity building in M&E, so that there will be a more widespread shift in the way M&E is understood and used.

Many of the changes to Benin's monitoring systems have grown out of various reforms to both the public financial system, and the adoption of a results-based management approach to public administration. For the past two decades, the ongoing LOLF reforms have modernised Benin's public finance management systems.[25] This shift towards better alignment between financial planning, accounting, and development

---

25   Before this important reform, the budgetary procedure was oriented towards means rather than results. However, since 2001, the LOLF has been instituting a performance-oriented budgeting approach. The budget nomenclature is broken down into major sets of public policies and these are then subdivided into programs. Credits are allocated to each public policy, as well as objectives, along with indicators to assess their achievement. This has paved the way for results-based planning at other levels of government.

results has led to better targeting of programme budgets, as well as stronger accountability for results. The introduction of program budgets across the public sector has been gradual and phased, beginning with six pilot ministries in 2000, before expanding across all ministries in 2006. This process of implementing results-based management was recognised as a strategy to implement appropriate development tools, and ensure citizens were being served by the public sector and has been critical to promoting a culture of efficiency, transparency, and accountability.

In Benin, the monitoring and evaluation systems operate quite separately and come together under the guise of 'results-based management' to produce knowledge for policymakers and politicians. Over the years, as evaluation has received increasing attention and support (especially from the donor community) the role and importance of the monitoring system has been somewhat subsumed into thinking of monitoring only as part of a larger evaluation system. Prior to the central role evaluation has played in Benin's development planning, there has been a long monitoring tradition across the public administration. Benin's membership in ECOWAS and UEMOA has meant that there are a range of directives, convergence criteria, and policy guidance in terms of the monitoring systems required for project and programme implementation. This entrenched bureaucratic environment has laid a foundation for building institutions equipped with monitoring capacity to enhance public sector performance.

Before the implementation of the National Evaluation Policy, certain routine M&E activities were carried out by the public administration, but they were done without a broader link to results-based management. When the government introduced reforms to implement the Paris Declaration (2005), along with donors working in Benin, they were carried out in line with a principle of mutual commitment to improve aid effectiveness. This included a practical and concrete framework aimed at improving the quality of aid and its impact on development. However, there is limited engagement with impact evaluation for many reasons, such as the weak national capacity in impact evaluation methodology, the mismatch between the evidence from impact evaluation and the politicians' agenda, and the high cost of impact evaluation.

Benin has been a 'rising star' in its push for developing evaluation capacity in recent years, but what impact has this had in terms of a monitoring system that underlies these evaluations? This chapter will detail the public sector structures in place for monitoring and consider the implications these have had for governance and decision-making in Benin's political and institutional context.

The chapter is structured as follows:

- Broad introduction to the institutionalisation of M&E to date in Benin.
- The country context of Benin, including specific monitoring projects and programmes within its public administration.

- The system that has been developed and put in place to assist with monitoring the activity of government projects and programmes.

- The role of monitoring and evaluation in promoting good governance at the country level.

- Lessons learned and good practices in the development of national monitoring and evaluation systems.

- Analytical discussion on the country perspective of Benin.

Context of Monitoring in Benin's Public Sector

Benin has two levels of government, with 22 national departments, and 77 local governments and/or municipalities. Historically, public policies and programmes in Benin have not been evidence-based and there has been no culture of accountability, except to fulfil conditionalities imposed by development partners (Goldman et al. 2018). A diagnostic study carried out in 2010 by Davies and Houinsa found that evidence use, from both monitoring and evaluation, had not been institutionalised in the process of planning or implementing programmes, projects, and policies in Benin, with those mandated to collect and use knowledge working in silos.

New reforms taken in 2016 to strengthen good governance focused on improvement in the coordination of government action, moving the mandate for this coordination to within the presidency, with support from the Ministry of Development. This meant that the key office for monitoring the performance of ministries and municipalities to improve service delivery, the Bureau of the Public Policies Evaluation, was now situated close to the highest level of decision-making.

Having those responsible for monitoring and coordinating government action in the central office (Presidency) allowed for a stronger focus on supporting the National Development Plan (NDP). This plan establishes indicators to be monitored, and prompts the development of standards, norms, and regulations for the monitoring of government policies and programmes. While institutionalising this system remains a work in progress, the creation of institutional infrastructure now means that ministerial plans are required, and provide the information required for decision making.

In addition to the NDP, budgetary reforms have placed increasing importance on monitoring, thus strengthening the monitoring function in the context of the implementation of the budget reform and the adoption of results-based management. This has further led to growing interest in the development of evaluation and its management by the government, which resulted in its institutionalisation to strengthen RBM and to formalise the evaluation of public service through the creation of a dedicated ministry.

In recent years, the role of the Bureau of Public Policy Evaluation has expanded to include monitoring public participation. This has happened because of the establishment of a transversal performance indicator for each department around a specific form of public participation. The goal is to expand the monitoring and evaluation function to be a visible tool to promote transparency, strengthen trust in government, and strengthen systems of governance at a national and local level, while remaining in line with the institutional reforms and calls from the public to improve governance and transparency.

National Monitoring System Structure

Monitoring systems within Benin's public service are found in:

- The Monitoring and Evaluation architecture (to strengthen results-based management related to the NDP).

- Budget monitoring.

- Program planning and implementation at both the national and local levels.

Table 9.1: Outline of monitoring system components and actors

Actor	Monitoring Responsibilities	Link to other areas	Actions to improve their monitoring function
WAMEU	Directives implemented through NES	Monitoring is part of NES – system for accountability	LOLF reforms
Parliament	Receive performance reports, oversee budget decisions at the medium-term level		Have received training on better oversight
BEPP			
M&E units	All ministries and depts - responsible for monitoring performance, quarterly report on annual performance plans, as well as biannual reports to Cabinet. Over 80% of reports reflect progress against agreed on indicators which managers need for planning		

Since the establishment of the BEPPAG in 2008, there has been significant political will to ensure that monitoring and evaluation become a strategic management tool for development. This demand has been deliberately cultivated through the establishment of M&E units in national departments and municipalities. Figure 1 below illustrates the actors involved in the national evaluation system. This has

guided the stakeholder engagement for evaluation in Benin, given the need for widespread capacity building across all levels of government.
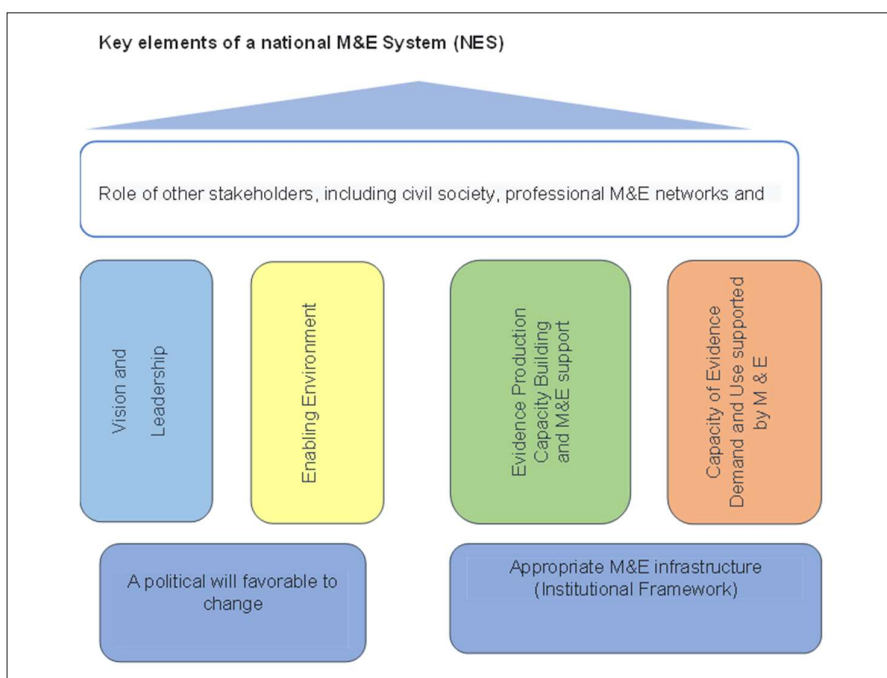
Figure 9.1: *Adapted by the author from Lahey, 2013. Framework for a National M&E System.*

Monitoring as Part of the National Evaluation System

The Government of Benin has recognised that monitoring is a critical function of evaluation, as monitoring feeds evaluations and influences decision-making. In Benin, the M&E systems have been built-in directives from WAEMU, and monitoring is seen as an important part of its National Evaluation System (NES) that leads to accountability. Current reporting systems allow the government to review programmatic progress on an annual basis and answer not only to Parliament but also the public on the results achieved, resources allocated, and policy decisions made as a result of implementation experience.

Monitoring and evaluation are two separate concepts. However, there are some commonalities. Monitoring and evaluation are both geared towards learning from what you do and how you do it, with a focus on cost-effectiveness. Monitoring also follows the criteria set out by the Organisation for Economic Cooperation and Development criteria (OECD) such as efficiency, sustainability, and impact.

They have similar purposes in that they provide information that can help inform decisions to be made, improve performance, and achieve set results when used to adapt programme performance.

In 2012, a national evaluation policy was adopted to improve the effectiveness and results of government programmes, governance practice, and governance decision-making, including an institutional framework established to define the mechanisms for conducting evaluations. This framework included guidance on selecting evaluations, the engagement of stakeholders, the dissemination of results, and the monitoring of the implementation of recommendations. To assist with impartiality, independent service providers undertook the evaluations, whether universities, consultants, or civil society organisations. However, since the focus of the NES is on evaluation, there has been very little capacity development or work done to build strong monitoring systems or personnel.

All departments and ministries in Benin have M&E units, considering that M&E is not a function of programme managers. In fact, the head of M&E units are at least as senior as the managers responsible for programme implementation. These M&E units are responsible for monitoring performance, providing quarterly reports on annual performance plans, as well as biannual reports to Cabinet. Over 80% of reports reflect progress against agreed on indicators which managers need for planning, in an attempt to ensure a linkage between monitoring reports and evidence usage.

Monitoring as a Learning Path

Monitoring helps managers and policymakers by giving them evidence to see what results have been achieved based on inputs made, including budgeting and human resources (Goldman et al. 2018). Benin has deliberately aligned its M&E process to both the sustainable development goals (SDGs) and voluntary national reviews (VNRs). This has guided the monitoring system in its priorities to strengthen coherence and cooperation, with a singular goal of improving data production for learning and better decision-making. Both quality and availability pose significant challenges to the production and use of data for policy-learning processes. The monitoring system straddles the issues of both production and use, with a goal of meeting both needs, overcoming cultural hurdles related to the perceived need for - and use of - monitoring data that continue to inhibit the use of monitoring for learning.

Monitoring Capacity

The institutionalisation of M&E in Benin has been taking place in tandem with the benchmarking of skills, to better understand the skills required to be responsible for an M&E function or portfolio at all levels. As the organisational needs and institutional context becomes clearer, human resourcing units within government will be able to better determine the various levels at which civil servants with M&E responsibilities

need to be appointed. Now, while a high level of education is required to be appointed to an M&E position, there are no other specific skill requirements.

It is an ongoing process to select and recruit appropriate personnel for monitoring units across the country. Furthermore, there is a requirement for training programmes for M&E officers. While there are some capacity-building activities offered, dedicated training in data collection and monitoring needs to be strengthened to ensure that those who assume monitoring (and evaluation) responsibilities can better understand the function and qualify to focus in this area. Unfortunately, there are relatively few professional courses available to either monitoring or evaluation practitioners in Benin.

Monitoring for Decision Making: An Example of How Monitoring Contributes to Good Governance

Periodic reports on the progress of the Public Investment Program (PIP) represent the basis for assessing the achievement levels of development projects and programmes in Benin. The PIP is a key part of the M&E mechanism of the Growth Strategy for Poverty Reduction (SCRP 2011-2015), and this, along with quarterly reports and other monitoring outputs, help different government actors to take appropriate decisions, to provide appropriate solutions to the shortcomings observed, and to anticipate and engage citizens for the improvement of the development and implementation of the PIP (and thus, the Growth Strategy for Poverty). The development of these reports follows a methodological approach consisting of four essential steps, which are: (i) collecting information, (ii) data management, (iii) drafting, and (iv) validation.

The Public Investments Program (PIP) monitoring system is governed by an institutional framework centred on four levels that support the national performance monitoring system. The Ministry of Development, Economic Analysis, and Prospective (MDAEP), through the Directorate General for Projects and Programmes Monitoring (DGSPP), anchors the monitoring of projects and programmes at the national level.

The mechanisms and tools used at each level to produce evidence for decision-making are outlined below:

1. Local Level – Project Management Units (PMUs) and Town Halls.

   The main role of PMUs is to ensure effective and efficient execution of projects, and to monitor the progress of government work. These PMUs periodically report on the performance of their activities to the Directorate of Planning and Prospective (DPP) of their line ministries where evaluation units are located. Some of these projects, given their size, have monitoring and evaluation units that facilitate this reporting work. The town halls ensure the proper execution of the projects by the central state. They produce monthly physical monitoring reports, which they send to departments (see below).

2. **Regional Level – Prefectures and Departmental Directorates (DD) of State ministries**.

    The DDs act as relays for the Directorate of Planning and Prospective DPPs at the regional level. They monitor projects carried out by departments. Reports from town halls in all sectors will be sent to the prefectures, which in turn will send it to the Departmental Directorates of Development and Prospective (DDPD). These reports are studied during the Departmental Administrative Conferences (CAD), which is a consultation framework within which views on the development of work in the departments are harmonised. At the end of each quarter, the DD sends the summary of their respective sector to their 27 DPPs, focused only on achievement of outputs. DDPDs, in turn, summarise the department reports from all sectors. Coordination of the CAD to meet on the right date before the summaries reach the central level, either the DPP or the DGSPP, is of utmost importance to streamline monitoring and planning.

3. **Sectoral Level – Structures Represented by the DPPs.**

    These are responsible for centralising, synthesising, and validating the financial and physical data resulting from the execution of projects and programmes of departmental directorates. This data is used to compile a variety of documents, including Quarterly Review Reports, Departmental PIP Quarterly and Annual Progress Reports, and the Year-End Performance Report.

4. **National Level – The Directorate General for Projects and Programmes Monitoring (DGSPP).**

    This is responsible for centralising, analysing, and processing, in collaboration with the other coordination structures (DGIFD, DGB, CAA), the data sent by the DPPs from the ministries. The products resulting from this monitoring are the quarterly and annual progress reports of the national PIP, the dashboard for monitoring PAP indicators and the tour reports.

The PIP monitoring as described above is illustrated in Figure 2 below, illustrating the four levels of reporting on financial and physical monitoring for decision-making.
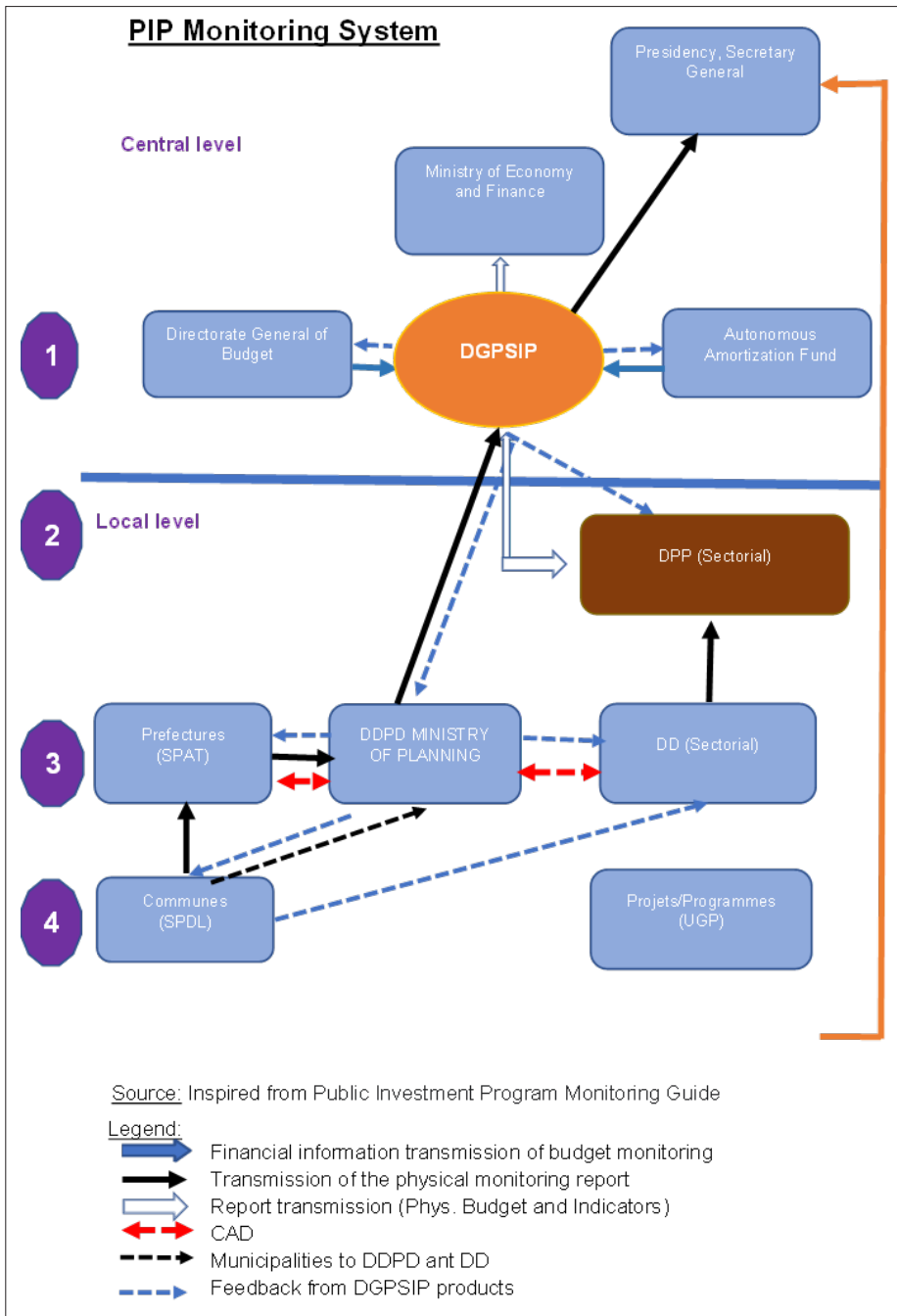
Figure 9.2:   *PIP Monitoring System*

Adapted by the author from the Public Investment Program Monitoring Guide.

For programming needs, four main systems have been developed to harmonise the budgetary and performance indicators monitoring of the PIP:

1. the System Harmonised Information on Public Investments (SHIIP);

2. the Integrated System of Analysis and Programming of Public Investments (SIAPIP);

3. the Integrated System of Public Finance Management (SIGFiP) and;

4. the Harmonised and Integrated System for Monitoring and Evaluation of Public Investment Projects and Programmes (SHISEPIP), which takes information from SHIIP into account.

These systems are intended to be effective tools for monitoring, management, and evaluation of public investment projects and programmes, as part of a broader results measurement framework.

With the current reform (2021 onwards), a Programme/Project Monitoring Unit has been created within each ministry, comprising three specialist roles: i) financial management, ii) public procurement, and iii) monitoring and evaluation. To rationalise the use of government resources and better streamline decision-making processes, those units dealing with finance, human resources, and project management will work transversally to support the programme and project coordinators in the ministries, with the intention of generating an integrated set of information that can be used to improve decision-making.

Monitoring for Accountability Within the PIP

To provide appropriate responses to the requirements of budget reform and ensure improved synergy of actions taken at the national level, the institutional and organisational frameworks of the ministries responsible for development and finance have been bolstered through seminars as well as capacity-building activities, with financial and technical support provided by country donors. The aim of this was the improvement of the availability and accessibility of data, as well as reporting mechanisms that would allow for accurate and timely accountability reporting.

The main products resulting from this mechanism at the national level are, amongst others, the BGE, the PIP, the BGE execution reports, quarterly progress reports, and the annual PIP.

Local Level Monitoring for Accountability

At the local level, Benin has opted for decentralisation to lay the foundations for a sustainable and local democracy to ensure grassroots development. Thus, the municipalities are obliged to generate and adopt their municipal development plans. The Communal Development Plan (PDC) expresses the priorities of municipal policy

and constitutes an instrument for framing short- and medium-term development actions. In the PDC preparation guide (2008), it is stated that "the PDC are aligned to the national development plan which itself is aligned to SDGs Agendas (United-Nations and African Union)".

Therefore, the PDC must fit into the development and performance management cycle at the national level. Indeed, the options defined in the PDC must fit into the national and regional development priorities that emerge from both national and sectoral reference documents. These include:

- Strategic Development Goals, Space Agenda.

- The Municipal Development Master Plan.

- National and sectoral policies, sector strategies, and the Growth Strategy for Poverty.

As part of the monitoring of the implementation of the Communal Development Plan, the municipalities are assisted by the decentralised structures of the State. The Prefectures and the DDPD provide the necessary advisory assistance. This forms part of the overall Program of Government Action, supported by the Bureau of Public Policy Evaluation and Government Action in the Presidency.

The primary product of the PDC monitoring mechanism is the Annual Investment Plan monitoring report. Unfortunately, not all the municipalities have the necessary skills for monitoring the projects and programmes included in the Annual Investment Plan. However, the support of the BEPPAG is critical to boost capacities through the M&E Units across the 77 municipalities.

Good Practice Examples from Benin's Monitoring System

The monitoring and evaluation system is an important tool to consolidate Benin's various monitoring systems for the implementation of the policies, programmes, and projects. Collaborative partners outside of government are essential to the development of specific tools linked to the monitoring of PC2D and government action and assist by augmenting technical capacity.

The strategy adopted by the Bureau of Public Policy Evaluation has centred on benchmarking M&E practice. This strategy has aided in the integration of this practice into international evaluation networks to demonstrate the value of the concept to diverse stakeholders, while also addressing issues of capacity, and being led by good practice. Having a dedicated structure at the national level that centralises both monitoring and evaluation is valuable as a strategy that strengthens evidence use in government. The Bureau of Public Policy Evaluation supports the public service with capacity building activities and standards definitions, while providing tools and procedures for improving results-based management.

Key Success Factors

The recent strengthening of the M&E institutional framework has led to substantial improvements in decision-making processes and the delivery of public services. Strengthened reporting mechanisms harmonised from the programme and projects M&E Units to the Presidency of the Republic through the M&E Units in the ministries and directorates located in the prefectures has been a key success. This institutional chain strongly contributes to strengthening the practice of accountability at both the local and national levels.

The following few elements have facilitated the culture of performance monitoring in the country:

- Political will at the top of the state supporting the institutionalisation of monitoring evaluation.

- Directives form WAEMU (RBM and LOLF).

- National institutional anchoring (Ministry of Development and Government Action Coordination and Ministry of Economy and Finance).

- Leadership to a dedicated national structure.

- Strong and dynamic partnerships with multinational and bilateral organisations: Twende Mbele and WACIE, etc.

- Effective regional and international networking and learning from other countries in Africa.

Constraints

In practice, the existence of an M&E system has not guaranteed the production and use of monitoring data, nor fed the multitude of decision-making processes. Without the above-mentioned institutional factors, it does not automatically promote a culture of good governance practices. Other considerations such as global country awareness, political will, decision-makers' strong engagement, and country capacities in monitoring and performance reporting mechanisms are determinants. Performance culture in Benin meets some facts can be strengthened by addressing the following factors:

- The lack of a legal framework to support the process (which has been addressed through the recent reforms).

- Poor systematisation of data collection mechanism using monitoring (the current digitalisation will improve this).

- Weak national capacity ensures the efficiency of the system which limits the synergy between monitoring and evaluation. To address this, both aspects of the M&E

function were taken over by the current Ministry of Development and Coordination of Government Action, since July 2022.

- The weakness of national financial resources allocated to monitoring (or evaluation) activities and quality data systems.

Key Debates in Monitoring Systems: Perspectives for Benin

Benin's approach to the development of its M&E system has enabled it to build its international reputation and to the innovative cooperation programmes presented above with its partners. The sharing of experience and learning from other country governments and technical partners has led to the development of collaborative governance tools that are improving the quality of transparency in public management in Benin. However, essential questions remain unanswered regarding the improvement of this system, which may allow it to optimise support of the production at the national level of quality of data evidence that can be used by all development actors.

Central Level

The main concern at this level is to digitise monitoring systems such that data on the implementation of public action is generated in real-time to enable accurate decision-making. In fact, since the COVID-19 pandemic, the government has undertaken a vast digitalisation reform within the administration, including the Councils of Ministers and online administrative meetings. Procedures, mechanisms, and working tools are being digitised in this context.

The feedback of digitised information in such an environment will improve the speed of data processing, increase transparency with citizens, and strengthen the overall quality of the data. This will be done by developing statistical capacity in line ministries to produce data related to their sector public policy.

Local Level

The creation of further synergies between municipalities and the central level to harmonise the collection of data and production of information that can be used for decision-making in programming and budgeting is an area of need. Current efforts at streamlining and institutionalising monitoring systems to provide data and evidence on the implementation of communal development plans are underway and are yielding promising results, albeit with a need for much greater capacity.

Role Parliament

The issue of parliament's involvement in monitoring and evaluation activities remains a challenge in national monitoring and evaluation systems, due to the need

to respect the separation of powers in a democratic regime. Indeed, the separation of powers within a strong presidential democratic regime like Benin means that government initiatives towards parliamentary monitoring remain ceremonial and bureaucratic. Parliament has remained stuck in the traditional way of taking charge of its prerogatives by using the conventional tools of oral or written questions to control government action.

In addition, the parliament is not currently properly equipped to conduct M&E activities and positions itself as a consumer of M&E activities rather than as an active actor in the production or oversight of monitoring. Even at this level, its capacities in this area are limited, which means that great advantage can be taken of it to control government action. Although an M&E Unit has been established inside the parliament administration, it struggles to achieve results to support parliamentarians with their M&E issues. Somewhat outside of its mandate, the critical role of parliament in oversight of public action has meant that BEPPAG, with the support of UNDP, has begun supporting the M&E unit in parliament with capacity building activities.

In recent years, BEPPAG, in partnership with the Beninese branch of the African Parliamentarians' Network on Development Evaluation (APNODE), has undertaken awareness-raising activities, capacity building, and some specific training for Beninese parliamentarians as to how M&E fits into their mandate and can aid in their duties. Thus, more and more parliaments are using M&E data, not only to understand the processes of implementing government action, but also to better exercise their control function for better governance.

Conclusion

Development policies and programmes at the global level demand more performance and efficiency, not only from the government but also from partners and donors. Monitoring is a mechanism to report on the commitments made to ensure good governance. It is a strategic steering tool that aids in the informing of decision-makers, both as a function of evaluation for evidence production and as a role of accountability.

Therefore, the actors from the national monitoring system institutional framework must take ownership of it according to their administrative context and pool their experiences to inform the decision-making with high quality administrative and statistical data to improve the governance.

This chapter calls for the establishment of solid engagement of the Directorate for Evaluation for the strengthening of the systems and the law on evaluation to create real effectiveness across the country, so that the gains of recent reforms can be sustained and scaled.

Reference List

Amouzou, A., Kanté, A., Koffi, A., Maïga, A., Munos, M., Walker, N. & Winch, P. 2020. Regional Scoping Study for the West Africa Capacity Building and Impact Evaluation (WACIE) Program.

Dumitrescu, A., Houndolo, D., Sossou, D. & Mama, B.B. 2019. Synthèse de la littérature sur l'évaluation d'impact dans l'UEMOA. p. 2.

Arnold, R.D. & Wade, J.P. 2015. A Definition of Systems Thinking: A Systems Approach. Steven's Institute, New Jersey. https://doi.org/10.1016/j.procs.2015.03.050

BEEPPAAG. 2017. Guide Méthodologique National d'Évaluation. UNICEF.

Burnes, B. 2004. Kurt Lewin and the Planned Approach to Change: A Re-appraisal, Journal of Management Studies 41:6 September 2004, 0022-2380. https://doi.org/10.1111/j.1467-6486.2004.00463.x

Benin. 1999. LAW N˚ 97 028; 15 January 1999: Organisation of the territorial administration in the Republic of Benin.

Chirau, T., Waller, C., & Blaser-Mapitsa, C. 2018. The National Evaluation Policy landscape in Africa. A Comparison. Johannesburg: CLEAR-AA and Twendé Mbele. Policy Brief.

CLEAR-AA. 2018. Mere compliance or learning – M&E culture in the public service of Benin, Uganda, and South Africa. p.3.

Davies, I. & Houinsa, D. G., 2010. Étude diagnostique des capacités nationales en évaluation au Bénin.

DPSIP. 2013. Guide de Suivi du Programme d'Investissements Publics.

Goldman, I., Byamugisha, A., Gounou, A., Smith, L.R., Ntakumba, S., Lubanga, T., Sossou, D. & Rot- Munstermann, K. 2018. The emergence of government evaluation systems in Africa: The case of Benin, Uganda and South Africa. *African Evaluation Journal 6(1)*, pp. 1-11. https://doi.org/10.4102/aej.v6i1.253

Government of Benin. Decree No. 2021 of July 28, 2021, Setting the standard structure of ministries.

Chapter 10: Frontline Service Monitoring Systems as Catalysts for Improved Service Delivery in South Africa

Nedson Pophiwa and Caroline Nuga Deliwe

Introduction

The South African government introduced monitoring of frontline service delivery almost a decade ago in response to increasing dissatisfaction and alarming levels of service delivery protests (Cheruiyot, Wray, and Katumba 2015:27). The Zuma administration, which assumed office in 2009, recognised the significance of adopting an outcomes-driven approach to governance due to citizens' demand for change in their lives. A few months after his election as president, Jacob Zuma visited communities in the town of Balfour (under Dipaleseng Municipality in Mpumalanga province), where violent protests erupted over poor service delivery and community outrage regarding boundary demarcation. The visit revealed that Balfour residents were dissatisfied with the demarcation of their municipality from Gauteng to the Mpumalanga province, and they also expressed grievances about the lack of medical personnel at the local clinic. Additionally, there was no police station, a housing backlog existed, and there was a lack of proper sports and recreational facilities (The Presidency 2010). Moreover, the municipal mayor was absent from the office during the visit, citing a stomach problem. Following the surprise visit, the president established a ministerial task team to address the community's grievances and provide regular reports to the community. Zuma conducted other surprise monitoring visits during his first year in office, laying the groundwork for an institutionalised approach to monitoring frontline service delivery, which later became one of the mandates of the Department of Planning, Monitoring, and Evaluation (DPME).

In this chapter, we will profile the evolution of frontline monitoring systems in the public sector. At that time, the government faced a significant challenge in improving effectiveness, efficiency, accountability, and transparency. The government aimed to

invest greater focus on public sector performance to enhance service delivery for all citizens (Plangemann 2016:73). The African National Congress (ANC) led government recognised the importance of establishing a comprehensive government-wide system to transform the performance culture burdened by apartheid legacies. To achieve this, the government adopted several policies and approaches. These included introducing an outcomes approach guided by 14 government priority outcomes, implementing the Management Performance Assessment Tool (MPAT) to assess the quality of management performance in national and provincial departments, creating a similar version for municipalities, establishing a system to monitor front-line services, and developing a comprehensive national evaluation system that promotes the use of research evidence (Plangemann 2016:72).

This chapter focuses on the frontline monitoring systems of DPME, which consists of programmes designed to monitor the quality-of-service delivery to users of government services. The main emphasis is to demonstrate the usefulness of monitoring in enhancing the quality of frontline services. DPMEs frontline monitoring systems include the Presidential Hotline (PH), Frontline Service Delivery Monitoring (FSDM), Citizen-based Monitoring (CBM), and Special Project/Izimbizo (SP). These systems were strategically established to complement and verify other routine monitoring systems in DPME and the government, while also addressing service delivery challenges (DPME 2019). The adoption of this service delivery monitoring approach, according to DPME (2018), marked the beginning of a process to improve government performance, support service delivery plans, and strengthen intergovernmental relations.

Given limitations in data availability, this case study will primarily focus on the frontline service delivery monitoring (FSDM) launched in 2011. The argument put forth is that while the introduction of FSDM provided public officials with evidence to enhance the quality of frontline services, its capacity to ensure the implementation of its findings was limited. As demonstrated in this chapter, participating public offices were able to enhance citizens' experience with services by monitoring and improving simple aspects such as visible signage or reduced waiting times. Nonetheless, it remains crucial for public officials to foster a culture of monitoring service provision to identify challenges and address them promptly, thereby preventing citizen unrest.

Monitoring – An Analytical Framework

The literature has varying interpretations of what monitoring entails. The word monitoring comes from the Latin word *monere*, meaning "to warn" (Kettner, Moroney and Martin 2013).

Monitoring serves as a feedback mechanism, alerting a human service administrator when a programme's implementation deviates from its original design. This enables

the administrator to take corrective measures and realign the programme with its intended design. The primary purpose of monitoring is effective programme management (Kettner, Moroney, and Martin 2013).

For programmes that have transitioned from the development stage to active operation, programme process assessments fulfil management requirements by providing information on service delivery and coverage. This includes evaluating the extent to which a programme reaches its intended target population. Additionally, such assessments may offer insights into participants' reactions and experiences with the programme (Rossi, Lipsey and Freeman 2019)

In comparative terms, the monitoring and evaluation literature tends to pay less attention to monitoring than it does evaluation. Empirical studies which provide case studies of monitoring of programme implementation, or which monitor progress in government service provision to citizens, have room for development as processes for learning. Porter and Goldman (2013:6) provide a simple yet illustrative definition of monitoring:

> "Monitoring is a management function focused on tracking if you are doing what you intended, whether at the programme level or for higher level national goals. Monitoring helps you to know how you are progressing compared with the plan, what is being produced, and what evaluative questions to ask. Monitoring data does not enable you to understand why something is happening. When evaluative conclusions are drawn at the apex of government from monitoring evidence alone, there are likely to be errors: claiming an effect when there is none, claiming no effect when there is one, or a lack of understanding of what is causing what."

The latter part of the definition of monitoring is crucial as it emphasises the importance of staying within its intended scope. In less developed countries, particularly in Africa, monitoring tends to hold a more prominent position than evaluation. While there has been an increase in commissioned evaluations in Africa in recent years (some of which are catalogued in the African Evaluation Database), monitoring still takes precedence in practice, occurring on a larger scale.

This chapter does not aim to lament the dominance of monitoring in certain sectors of the public service. Instead, it seeks to reflect on the positive contributions that monitoring makes. Adjusting to the new normal has become essential for policymakers to ensure that monitoring is not negatively impacted by the need to contain the spread of COVID-19. Through the COVID-19 pandemic, numerous initiatives have emerged that focus on continuous monitoring of dynamic programmes using mobile phones and short surveys. An example of this is the World Bank's "High-frequency monitoring of COVID-19 impacts project", which employs phone surveys to monitor the effects of the pandemic on households and individuals in over 100 developing

countries across all regions (Fu and Sanchez-Paramo 2020). The pandemic has underscored the significance of evidence-based decision-making, and data derived from monitoring plays a crucial role in informing such decisions.

Monitoring of public services is a crucial strategy to ensure that citizens receive the necessary services from the government. In areas where absenteeism among frontline public servants, such as teachers and healthcare workers, is prevalent, monitoring becomes an essential approach. A survey conducted in Bangladesh, Ecuador, India, Indonesia, Peru, and Uganda by enumerators who made unannounced visits to primary schools and health clinics revealed that absenteeism rates ranged from 19% among teachers to 35% among health workers (Chaudhury, Hammer, Kremer, Muralidharan, and Rogers 2006:91). While the study focused on the presence of public servants in their workplaces, a significant proportion were physically present but not fulfilling their assigned duties. Moreover, the study indicated that absenteeism was widespread across different hierarchies, with doctors and school principals being more absent than lower-ranking workers. Additionally, men were found to be more frequently absent than women (Chaudhury et al. 2006:92). A subsequent follow-up survey reported a lower teacher absence rate, suggesting that monitoring could have an impact on reducing absenteeism (Chaudhury et al. 2006:92).

This idea of monitoring as an effective means of addressing absenteeism is supported by Muralidharan, Das, Holla, and Hopal (2017:117), who argued that "reducing teacher absences by increasing school monitoring could be over ten times more cost-effective in reducing the effective student-teacher ratio (net of teacher absence)". Considering the situation in many developing countries, it becomes crucial to understand the underlying causes of absenteeism and other forms of poor service delivery that hinder governments' ability to provide essential services to their citizens.

Government departments and agencies adopt a two-pronged approach to monitoring. The first approach is sector monitoring, which involves monitoring the performance of the sectors for which they are responsible in delivering services or providing regulatory oversight.[26] The second approach is programme monitoring, which focuses on monitoring the performance of specific programmes and policies that they are implementing. In this chapter, we will argue that shifts in incentives have played a crucial role in shaping the necessity and focus of monitoring, based

---

26   The purpose of the DPME's sectoral monitoring programme is to develop the country's long-term vision and national strategic plans and contribute towards better outcomes in government through better planning, better long-term plans, greater policy coherence, and a clear articulation of long-term aspiration. The main responsibility of the programme is to institutionalise and strengthen planning in government by facilitating the development of sectoral plans, ensuring coherence between plans, policies and service delivery across government, ensuring high-level priorities are fed through into plans across all spheres of government and engaging stakeholders on the output of the planning process to ensure buy-in. https://pmg.org.za/tabled-committee-report/4497/

on a comparative analysis of selected cases of public policies and programmes in post-apartheid South Africa. Specifically, the chapter will analyse the monitoring of frontline service delivery, which is a collaborative programme between DPME and the nine Provincial Offices of the Premier.

Impetus: Service Delivery Challenges and the Need for a Government that Listens

In this subsection, a discussion of the main drivers of frontline services monitoring will be provided. These drivers provided an impetus for the kinds of monitoring which DPME introduced in the public sector since its inception. The discussion will begin with unpacking the levels of citizens' discontent with poor service delivery and how they triggered the response by the government.

Citizens' Discontent with Poor Service Delivery

The introduction highlighted that the Zuma administration took office in 2009 amidst widespread dissatisfaction with service delivery. The initial enthusiasm following the end of apartheid had diminished, and citizens had grown increasingly frustrated with unfulfilled promises. A study conducted by the Department of Public Service and Administration revealed a significant decline in citizen satisfaction with service delivery during the final years of Thabo Mbeki's presidency, dropping from 75% in 2006 to 58% in 2008.

In 2008 and 2009, service delivery protests had become commonplace, spreading across townships in various parts of the country. During the first half of 2009 alone, 26 public demonstrations, some of which turned violent, were recorded in seven of South Africa's nine provinces. A parliamentary study found that the primary grievance of the protesters was poor service delivery, particularly concerning water, electricity, sanitation, and waste removal.

Despite the challenges, Zuma was re-elected as the leader of the ANC in 2012 and as the president of South Africa in 2014, securing slightly reduced support with 62% of the vote (Centre for Public Impact 2016). Research conducted during that time identified three main deficiencies within the South African government that contributed to poor service delivery: a lack of accountability at the ministerial and upper managerial level, inadequate planning in some ministries with a failure to align activities with departmental plans, and ineffective coordination between ministries in formulating and implementing policies (Friedman 2014).

The primary stakeholder in the Government Performance initiative was the DPME, which played a central role in its implementation. The president himself was deeply involved and had prioritised the establishment of the DPME as one of his initial official

actions. Other government ministries were also key stakeholders, with some fully embracing the new approach and actively participating. Examples include the health and education ministries, as mentioned in the previous discussion on public impact.

The Treasury was another engaged stakeholder, as they had already contributed significantly to fostering accountability in the civil service. They had implemented a Programme Performance Information reporting system that required ministries to report data that linked financial inputs to actual outputs. The Department of Public Service and Administration also played a crucial role, emphasising the importance of performance.

Furthermore, the Public Service Commission, an independent agency, published respected reports that evaluated ministerial performance based on nine criteria outlined in South Africa's constitution (Centre for Public Impact 2016). These various stakeholders collaborated to promote and implement the Government Performance initiative.

The Policy Challenge: Weak Monitoring and Accountability Mechanisms

The transition to democracy in South Africa coincided with global shifts in the 1990s that prompted governments worldwide, both in developed and developing countries, to reassess their service delivery methods (UNECA 2010). This was driven, in part, by the rising costs of social and infrastructure programmes that had characterised that era. As a result, there were calls for a reinvention of government, leading to the emergence of New Public Management (NPM) approaches and practices, which placed a strong emphasis on performance management, target-setting, and the measurement of outputs and costs (Hood 1991).

The NPM approach sought to minimise government involvement in service delivery, promote professional management principles and standards in the public sector, enhance financial reporting and accountability for managers, and incorporate market-led performance management and development strategies from the business world into the public sector. The United Kingdom and the United States were early adopters of these managerial approaches, and their influence spread globally, including through the conditions attached to official development assistance provided to developing countries, including those in sub-Saharan Africa (Hood 1991:4-5, Osborne and Gaebler 1992 and Power 1997).

South Africa, eager to reform and modernise its public sector, embraced these reforms as they resonated with the democratic changes introduced in the country in 1994. The country's new "fiscal" constitution placed a strong emphasis on financial management, complementing the focus on democratic and citizen-centric accountability policies in the post-1994 years (RSA 1996). Kuye and Ajam (2012) observed the unusually detailed specification of institutional arrangements for providing socioeconomic rights, including the right to basic education.

The *White Paper on the Transformation of Public Service Delivery* exemplified the policy position highlighting responsive public service delivery, and citizen-focused accountability in the context of a new vision of a working, social accountable public service:

> "The Public Service is also seen as still operating within over-centralised, hierarchical, and rule-bound systems inherited from the previous dispensation, which make it difficult to hold individuals to account because: decision-making is diffused; they are focused on inputs rather than on outcomes; they do not encourage value for money; they do not reward innovation and creativity; they reward uniformity above effectiveness and responsiveness; and they encourage inward-looking, inflexible attitudes which are at odds with the vision of a public service whose highest aim is service to the people" (RSA 1997:12, s1.2.9).

Despite the adoption of financial and performance management tools and approaches in the new democracy of South Africa, there was a lack of enthusiasm for market-based reforms or a diminished role of government in public service delivery. This sentiment was particularly pronounced among the new government, which viewed itself as a crucial driver of development (Fraser-Moleketi 2006:85 and NPC 2012:409). The Minister of Public Service and Administration in South Africa from 1998 to 2008 highlighted this discrepancy in her Master's dissertation. The ideological view of the government clashed with the more controversial aspects of new public management reform, creating tensions within the reform process.

> "South Africa needed to modernise traditional administrative practices hence the influence of NPM could be felt in certain areas of South African reform, notably the micro-level...based on the philosophy of 'putting the citizens first' which underpinned the movement in favour of 'citizens charters' in Britain in particular. By contrast not surprisingly the minimalist, neo-liberal ideology of NPM clashed with the democratic and radical approaches of the ANC especially with regard to the 'macro' sides of reform. But such association could not detract from the potential these tools offered to result in greater efficiencies in state administration which in turn could lead to improved service delivery and freeing up more money for infra-structural development and so forth -- all key aspirations of the ANC's transformation agenda" (Fraser-Moleketi 2006:62).

Before the establishment of the DPME, the responsibility for monitoring, including performance monitoring, was primarily held by limited entities such as the National Treasury, the Auditor General, and the Public Service Commission. According to the Constitution, the Auditor General and the Public Service Commission were mandated to independently monitor specific aspects of government and report their findings to Parliament. Additionally, three national departments possessed strong

legal powers related to regulation and, consequently, planning and monitoring: the National Treasury (in relation to departmental strategic plans, annual performance plans, and quarterly reporting), the Department of Public Service and Administration (regarding the performance of the public service), and the Department of Cooperative Governance (in terms of monitoring local government).

Before 2009, the government lacked a clear planning mandate and a national plan. The National Treasury managed the government's basic planning and monitoring system, which included five-year Strategic Plans (SPs) and Annual Performance Plans (APPs). However, the introduction of the outcomes system in 2009 posed a challenge of aligning departmental plans with the cross-cutting outcomes, as the SPs for 2009 - 2014 had already been defined. It took time to ensure proper alignment between departmental plans and the outcomes. In August 2014, the Medium-Term Strategic Framework (MTSF) was launched, providing an integrated framework for departments to align their 2015 - 2019 SPs/APPs. The Presidency assumed certain planning and monitoring roles, utilising the authority of its position and Cabinet decisions rather than relying on legal powers. The President also possesses constitutional powers to ensure efficient government (Goldman et al. 2012).

Indeed, efforts have been made to empower citizens and establish mechanisms for holding the government accountable. Constitutional democratic and oversight institutions, including Parliament, Chapter 9 institutions (such as the Human Rights Commission and the Public Protector), the Auditor-General, and the courts, play a crucial role in holding the government accountable for fulfilling its delivery mandate. These institutions are responsible for monitoring government actions, investigating misconduct or maladministration, and ensuring compliance with constitutional principles. Furthermore, ordinary citizens, civil society organisations, and the media also have an important role in holding the government accountable for delivering quality public services. Gumede (2017) notes that:

> "Public service and elective representative accountability have plummeted. Empowering ordinary citizens, civil society, communities, and the media to monitor government delivery of public services is more cost-effective, participatory and is likely to bring more accountability in government. It also increases democratic accountability, transparency, and quality of government. It will boost the capacity of the government to fulfil its mandate to serve the citizenry and increase the quality and sustainability of public services. Public service could become more responsive and reduce corruption more effectively" (Gumede 2017).

Despite citizens having voiced their concerns over poor service delivery, they seemed to lack the capability to then monitor the services as part of CBM.

Incentives for Monitoring Service Delivery

The chapter acknowledges that monitoring has evolved over the years from being preoccupied with monitoring inputs to monitoring outputs and, more recently, monitoring outcomes. In the early years after 1994, the need to report and demonstrate government performance largely incentivised the monitoring of inputs. For example, around service delivery, massive-scale housing projects monitored the utilisation of expenditure (inputs) and the number of houses built for the poor (outputs) to report the government's performance against its set targets to stakeholders. As Plangemann (2016:72) has observed:

> "The transition to democracy confirmed the vital role of the state in mediating social and economic relations in a highly unequal society. The public administration was geared towards a traditional approach, focussing on inputs and activities. There was limited inclusiveness of the policies, programmes, and projects to be designed, implemented, and monitored as well as limited diversity in the civil servants in charge of doing so. Overcoming the legacy of apartheid required providing opportunities for all citizens, by ensuring that government policies, programmes, and projects reached all intended citizens and had the intended impacts" (Plangemann 2016:72).

The move towards national development plans by the government has also incentivised monitoring to focus on outcomes. In recent years, public policies, state-funded projects, and programmes have been required to demonstrate the quality-of-service delivery and their impact on citizens' lives (outcomes). As a result, monitoring has evolved to prioritise improved service delivery quality, such as in the construction of houses, and ensuring sustainability. Plangemann (2016:76) argued that the new performance M&E system adopted by the government was intended to have a political effect and counter criticism from citizens and the international community, thereby helping to maintain and increase the government's power and popularity. However, the system was primarily designed to provide diagnostic and planning information rather than serving as an anti-corruption tool, which limited stakeholder buy-in.

The Response: Institutionalisation of Monitoring Service Delivery

Since the government that assumed power after the 2009 elections faced a number of pressures, including (i) persistent poverty and inequality; (ii) widespread service delivery protests at the municipal level; and (iii) loss of some political support in the 2009 elections, it necessitated the need for strategic direction. These pressures prompted the government to be frank about the poor quality of public services, corruption, and other governance problems, as well as to establish a political

consensus to enhance government performance, including through a greater focus on M&E. In 2009, the Presidency created the Ministry of Performance M&E, followed by the establishment of the Department of Performance M&E (DPME) in January 2010. Additionally, the Presidency formed the National Planning Commission (NPC) as an advisory body to develop a long-term 2030 plan (Goldman et al., 2012).

South Africa had already introduced the government-wide M&E framework in 2005 but the framework gained greater commitment from the government after the election in 2009. However, there were challenges pointed out in implementation:

> "While a strong point in the original conceptualisation of the GWM&E system was acknowledgement that the system was to be built over time, this approach has proved difficult in practice. With different paradigms of reform and views of the state in different agencies, this approach has led to central departments creating separate reporting systems. Similar information may be requested several times from departments, leading to additional reporting burdens on departments already battling considerable constraints in terms of skills and capacity, and suffering from reporting fatigue" (Philips et al. 2014:393).

In 2011, the government took the initiative to promulgate the National Evaluation Policy Framework (NEPF), which served as the foundation for a system of evaluations aimed at facilitating learning and improving the effectiveness and impact of government actions (IIED 2019). The NEPF established the groundwork for implementing evaluations throughout the government. A credible and expedited development of the national M&E system was achieved by integrating it into national planning processes at both the national and sub-national levels. This success can be attributed to a well-structured initial process that drew lessons from other national experiences and a favourable enabling environment driven by strong political commitment (IIED 2019). The national evaluation system (NES) is closely aligned with the national planning process, and the evidence derived from it directly informs policy decision-making. Monitoring has been effectively and sustainably introduced at the country level, providing valuable insights for decision-making processes. This accomplishment is a result of its emphasis on the government's strategic priorities and its seamless integration within the cycles of national planning and implementation (IIED 2019).

Initial evaluations were often constrained by a lack of data or a lack of good quality data. This meant that most of the initial evaluations considered efficiency and relevance, rather than effectiveness in achieving outcomes and impacts. The DPME has provided support, guidance, and training to address this challenge. Other historic reported challenges include:

- A culture of compliance, but a failure to M&E for the purposes of reflection and performance improvement.

- Duplication of reporting across departments.

- Not devising effective theories of change to underpin the M&E (IIED 2019).

The emphasis on outcomes meant that the "the new system lent greater coherence to national policy planning by requiring departments to organise their priorities and focus their efforts on outcomes that were determined by the presidency rather than on internally developed targets that often overlapped with other departments or left policy gaps." (CPI, 2016) The DPMEs most important contribution "was the use of data in formulating and assessing policies". However, the DPME's approach was not binding on government departments and although, "some ministries, including those dealing in health and education, embraced the system without pressure from the presidency, [others] did not embrace the outcomes approach" (CPI, 2016).

Table 10.1:   *Philips et al., M&E and planning roles of DPME. 2014*

M&E of national priorities	• Developing the MTSF/outcome plans (delivery agreements). • Monitoring (that is, tracking) progress against the delivery. • Agreements. • Evaluating to see how to improve programmes, policies, and plans. • Operation Phakisa – intensive planning, M&E, and problem-solving on priority programmes, building on the Malaysian experience.
Management performance M&E	• Assessing quality of management practices in individual departments (MPAT) at national/state level. • Assessing quality of management practices and delivery in local government (LGMIM).
M&E of frontline service delivery	• Monitoring of experience of citizens when obtaining services (joint with states) including citizen-based monitoring. • Presidential Hotline – including tracking responses and follow-up.

Government-wide M&E System	• National M&E policy frameworks.
	• M&E platforms across government – nationally, provincially.
	• Structures of M&E units/capacity development.
	• National Evaluation System.
	• Five-yearly reviews of changes in the country – for example, 20-year review.
	• Annual production of development indicators and the 20 years review are specific named documents.
	• Data quality issues.

A Description of the DPMEs Frontline Monitoring Systems

At the national level, the Framework for Strengthening Citizen-Government Partnerships for Frontline Service Delivery Monitoring was approved by Cabinet in August 2013 in response to the lack of citizens' experience of government services reflected within the government's monitoring systems, and the lack of a systematic use of this evidence to improve performance (DPME 2013). The framework summarises a range of citizen-based monitoring methodologies used both locally and internationally, with citizen-based monitoring pilot studies planned for 2014/15 (Cheruiyot, Wray and Katumba 2015:27).
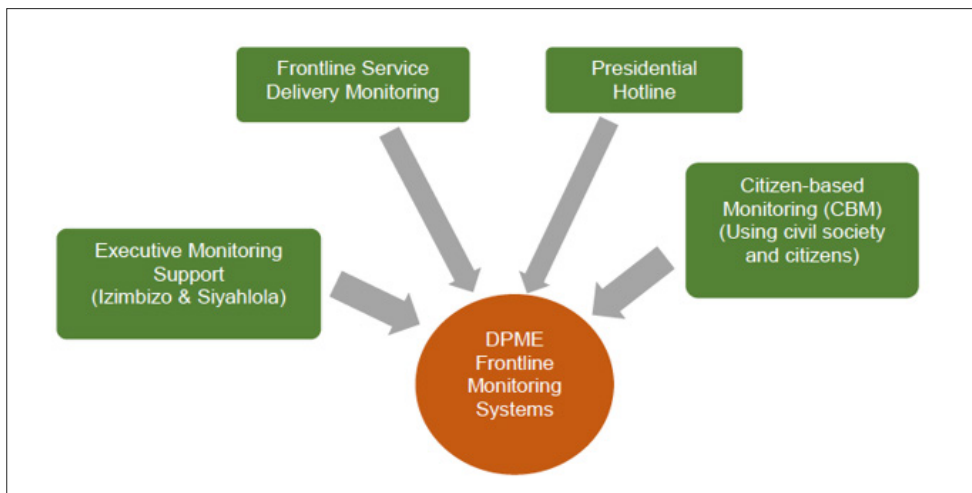

Figure 10.1:   *DPME, Frontline Monitoring Systems*

Presidential Hotline

In September 2009, former president Jacob Zuma took the initiative to establish the Presidential Hotline (PH), which provided a platform for members of the public to voice their concerns regarding the services they received from government departments and agencies. The establishment of the hotline was a response to the external pressures stemming from service delivery protests and widespread power cuts across the country. It demonstrated the President's commitment to addressing these issues as a top priority (Diga 2017:10). Between 2009 and 2013, over 180 000 cases were logged through the hotline, and the resolution rate saw a significant increase from 39% at its inception to 94% in 2013 (Graham 2015). In 2014, the then Presidential spokesperson, Mac Maharaj, emphasised that the hotline's goal was not only to achieve high resolution rates but also to provide a quality service to citizens (Maharaj, cited by Graham 2015). A survey conducted in 2015, which included feedback from 11 000 citizens, revealed that on average, 65% of respondents rated the hotline service as 'good-to-fair,' indicating that the government was making progress in this area (Graham 2015).

The key challenges of the PH were that the toll-free facility was costly, complex cases were reported, and there were unrealistic expectations from citizens. In the bid to overcome these challenges, the PH was being redesigned for two-way engagement through technological changes, in partnership with the Department of Science and Technology and the Centre for Science Innovation and Research.[27]

Citizen-Based Monitoring (CBM)

Citizen-based monitoring (CBM) aimed to strengthen the government's capacity to engage communities and citizens in monitoring service delivery and to ensure responsiveness to community experiences, expectations, perceptions, and needs. The DPME does not implement citizen-based monitoring directly but instead facilitates and enhances the capacity of government officials and departments. It also serves as an institutional repository for CBM methodology, good practices, and approaches. As a knowledge partner, the DPME provides support to government institutions involved in the implementation of citizen-based monitoring.[28]

The CBM programme actively supports the achievement of a Cabinet resolution (2013) that mandates all departments delivering public services to implement citizen-based monitoring. After an intensive two-year action learning process involving four service delivery departments and 34 government facilities (including police stations, health facilities, grants offices, and social welfare service points), DPME has transitioned

---

27   https://pmg.org.za/committee-meeting/28919/
28   https://pmg.org.za/committee-meeting/28919/

into a strategic support role. The aim is to promote the regular utilisation of citizen and frontline staff feedback and engagement to drive continuous improvement.

DPME's approach involves building the capacity of selected service delivery departments through hands-on support to officials. They assist in adapting and scaling DPME's three-step CBM model, which includes gathering feedback, utilising it to develop improvement commitments, and monitoring these commitments with the involvement of civil society and community structures. The objective is to develop a group of officials and civil society participants who actively contribute to the tools and knowledge required to utilise feedback and community participation, thereby fostering a capable and developmental state in collaboration with an engaged citizenry. Improving government responsiveness is the central goal.

Additionally, DPME plans to host ongoing discussions that bring together government and civil society to foster continuous dialogue on planning and monitoring matters.

Frontline Service Delivery Monitoring

Launched in 2011 by the DPME, the Frontline Service Delivery Monitoring (FSDM) programme monitors government facilities jointly with the Offices of the Premier across all nine provinces. DPME's roles and responsibilities include: (i) designing and maintaining the monitoring tools and the monitoring protocols; (ii) jointly conduct the monitoring visits with OTPs; and (iii) analyse findings and report to cabinet and national sector departments. Offices of the Premier (i) support the refinements of the monitoring tools and protocols; (ii) joint monitoring with DPME; (iii) Present findings to provincial HODs and MECs and other relevant forums; and (iv) monitor adherence to agreed improvement plans at provincial level (DPME 2014).[29] As stated in the 2014 FSDM Guidelines document, the FSDM Programme 'is not designed to cover all facilities in the public sector, but to demonstrate the value of on-site monitoring to selected facilities and catalyse service delivery improvement' (DPME 2014). It is argued that the FSDM programme rose as a response to a number of weaknesses in M&E in government, in which "problems are not treated as an opportunity for learning and improvement" and where "M&E is regarded as the job of the M&E unit and not all managers" (DPME 2015:8).

Improvements Monitoring and the Overall Aims of FSDM

Improvements monitoring is a key feature of FSDM. It is aimed at facilitating improvements in the performance of targeted frontline service delivery sites which have been performing poorly as observed during monitoring visits. Improvements monitoring focuses on identifying areas of weakness and to develop improvement

---

29   DPME (2014) FSDM Guideline 4.2.8

plans with facility, district, and provincial managers. Specific FSDM improvements objectives are:

- To institute improvements urgently in severe cases identified through the FSDM monitoring visits, as a means of strengthening service delivery;

- To monitor the implementation of improvement plans as developed by the line department together with DPME and OoP; and

- To monitor the improvements in scores from the initial baseline visit scores with the scores after the implementation of improvement plans.

An intended outcome of improved monitoring is the adoption of a culture of change in government towards increasing use of evidence in policy making, planning, and monitoring to inform improvements to plans and policies (DPME 2014). It remains the prerogative of line departments to implement the corrective measures emanating from improvements monitoring, despite developing plans with the assistance of the DPME (DPME 2014).

The purpose of the FSDM initiative is to strengthen the M&E practices of field-level managers and to ensure that decision-makers in head offices actively recognise on-site monitoring as a valuable source of evidence for decision-making. The aim is for decision-makers to utilise this evidence for prompt and decisive decision-making, as well as for driving systemic changes (DPME 2015:8). For successful implementation of FSDM, sector departments need to enhance planning and monitoring at the facility-level by establishing and continuously monitoring realistic norms and standards (DPME 2018:11).

Monitoring visits are a critical component of delivering FSDM. The objectives of these monitoring visits are: (i) to demonstrate to sector departments the value of onsite monitoring as a tool to verify the impact of service delivery improvement programmes; (ii) to demonstrate the value of obtaining the views of citizens during monitoring; to highlight successes and failures at service facility-level; and (iii) to support departments to use the findings for performance improvements (DPME 2015:8).

FSDM Approach and Methodology

The FSDM programme conducts targeted improvements monitoring whereby the selected sample of facilities is monitored every year to track improvements and regression, with a methodology that attempts to combine problem-solving facilitation and monitoring of results.

i. The Approach

The Improvements Monitoring approach consists of three activities:

1. DPME informs the national department (head office) senior management that a facility has been selected for improvements monitoring because of poor scores. The intention is for senior management to create an enabling and supportive environment in which facility-level managers can address the identified challenges.

2. A meeting is held at facility-level (led by DPME and OoP) to obtain progress with agreed improvements. The intention of this meeting is to facilitate acting on findings and to facilitate problem-solving between the different roleplayers.

3. The unannounced monitoring of improvements is conducted, applying the same scoring questionnaire tool used for the first assessment. A new score card is produced for the facility which reflects a longitudinal view of the scores, for each KPA, over time. A new score card is produced for the facility which reflects a longitudinal view of the scores, for each KPA, over time.

The monitoring team conducts unannounced visits to assess the quality of service delivery in frontline service facilities. During these visits, structured questionnaires guide interviews with citizens and staff to gather relevant information. The monitor also independently observes the adherence to the same standards assessed in the questionnaires administered to staff and citizens. Once the assessment is completed, a joint improvement plan is developed in collaboration with the facilities management and key decision-makers from the relevant department. The facilities are evaluated within the local context of communities, and the involvement of key stakeholders such as school governing bodies and business councils is crucial in the development of these improvement plans. The implementation of the improvement plan is then closely monitored. Key trends and findings are identified, analysed, and reported at the sector level and other decision-making platforms to address policy barriers that hinder effective service delivery (DPME 2018:11).

ii. Types of Facilities and Generic Performance Monitored

The FSDM programme monitors various types of facilities that represent the field offices of government service delivery. These facilities include Home Affairs Offices, SASSA offices, Police Stations, Health Facilities, Drivers' License Testing Centres (DLTC), Schools, Courts, Municipal Customer Care Centres (MCCC), and NYDA. The programme focuses on monitoring specific performance areas for ensuring the quality of service delivery in accordance with the policies and regulations set by the DPSA and the relevant national sector departments. These performance areas are as follows:

- Location and Accessibility

- Visibility and Signage

- Queue Management and Waiting Times

- Dignified Treatment

- Cleanliness and Comfort

- Safety

- Opening and Closing Times

- Complaints and Compliments Management

Figure 2 illustrates this focus on monitoring generic performance areas for quality service delivery (DPME 2018:11).
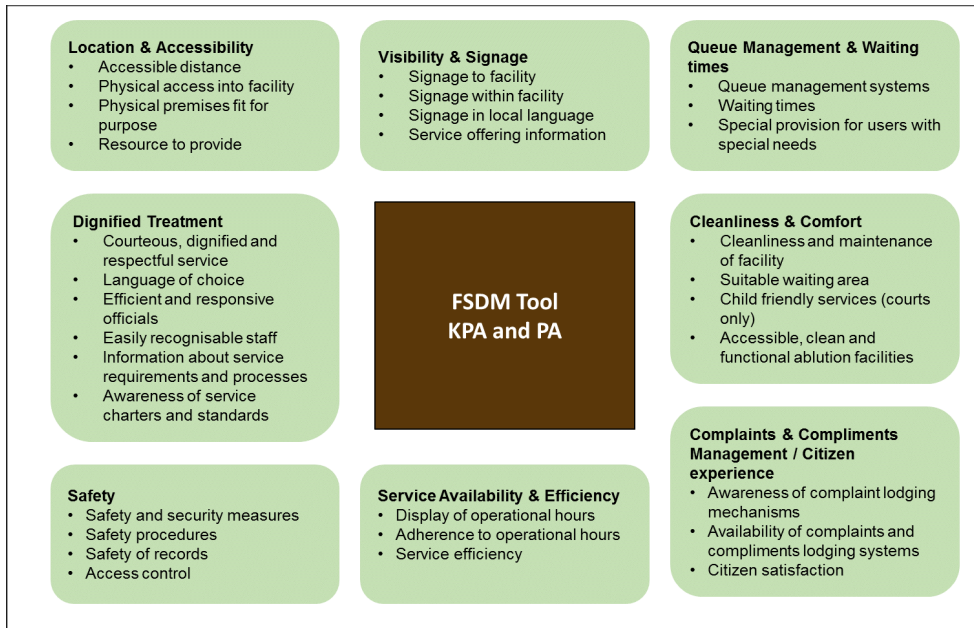

Location & Accessibility
- Accessible distance
- Physical access into facility
- Physical premises fit for purpose
- Resource to provide

Visibility & Signage
- Signage to facility
- Signage within facility
- Signage in local language
- Service offering information

Queue Management & Waiting times
- Queue management systems
- Waiting times
- Special provision for users with special needs

Dignified Treatment
- Courteous, dignified and respectful service
- Language of choice
- Efficient and responsive officials
- Easily recognisable staff
- Information about service requirements and processes
- Awareness of service charters and standards

FSDM Tool KPA and PA

Cleanliness & Comfort
- Cleanliness and maintenance of facility
- Suitable waiting area
- Child friendly services (courts only)
- Accessible, clean and functional ablution facilities

Safety
- Safety and security measures
- Safety procedures
- Safety of records
- Access control

Service Availability & Efficiency
- Display of operational hours
- Adherence to operational hours
- Service efficiency

Complaints & Compliments Management / Citizen experience
- Awareness of complaint lodging mechanisms
- Availability of complaints and compliments lodging systems
- Citizen satisfaction

Figure 10.2:   *DPME, FDSM Key Performance Areas, 2016*

Two types of monitoring are conducted: (i) baseline monitoring, which assesses the state of quality-of-service delivery, and (ii) improvement monitoring, which assesses year-on-year improvements. Baseline monitoring involves conducting an unannounced assessment (initial assessment) and holding a feedback meeting to present the findings to facility management. Improvement monitoring includes an

Improvements progress meeting to track progress against the improvement plans, followed by an unannounced re-scoring (re-assessment) to evaluate the changes made (DPME 2015).[30]

iii. Data Collection Tools

In terms of data collection tools, the FSDM programme utilises a structured questionnaire to gather data at the facility. The questionnaire is employed during both the initial visit and subsequent improvement monitoring visits. The questionnaire consists of three parts.

Part A focuses on the eight quality of service performance areas, with a set of questions directed towards each area. Ratings are assigned to each Key Performance Area (KPA) based on responses from three sources: facility users, staff, and the monitor. The ratings range from 1 to 4, where 1 represents poor, 2 signifies average, 3 indicates satisfaction, and 4 denotes above expectations.

Part B of the questionnaire involves rating the priority of improvement for each KPA on a scale of 1 to 4. All stakeholders—staff, citizens, and the monitor—are asked to rate the importance of improvement. For instance, if "safety" is deemed the highest priority for improvement, it will be assigned a rating of 1, whereas a rating of 4 indicates the lowest priority.

Part C of the questionnaire records observations pertaining to sector-specific issues. For example, at Police Stations, the effectiveness of the monitoring systems for police response time to calls for assistance is assessed. This aspect of the questionnaire may require further strengthening.

Lastly, Part D of the questionnaire captures the findings by formulating a draft action plan. It outlines the identified issues, specifies the necessary actions, assigns responsibilities to relevant parties, and sets deadlines for completion.

iv. The Facility Score-Card

The monitoring team summarises and documents the findings obtained from observations and interviews conducted by monitors with citizens (users) and staff. These findings are consolidated into a comprehensive summary report. The summary report includes the following components: (i) scores for each performance area derived from the inputs of three sources: citizens, staff, and monitors; (ii) prioritisation scores assigned by the three sources to indicate the level of importance for improvement; (iii) improvement plans outlining the necessary actions to address identified issue; and (iv) photographs that serve as visual evidence to illustrate the findings.

---

30   DPME 2015. FSDM Annual Findings Report 2014-2015.pdf

To complete the summary report, the monitoring team convenes and discusses the findings with facility management. This discussion, known as a "feedback visit," allows for mutual understanding and agreement on the action plan. It is important to note that the monitoring of a facility encompasses two essential activities, both of which must be completed before considering the monitoring visit concluded: the initial monitoring visit and the subsequent feedback meeting.

Performance in the first year:

- Location and Accessibility: 28 out of the 135 sites rated as poor (1), requiring intervention and 14 facilities rated as (4), being good practice.

- Visibility and Signage: 34 out of the 135 sites monitored were rated as poor (1- requiring intervention) and 8 sites demonstrated good practice.

- Queue Management and Waiting Times: 36 sites scored as needing interventions and 8 sites which demonstrated good practice.

- Cleanliness, Comfort, and Safety: below acceptable conditions found in 49 facilities, with some good practices in 9 facilities.

- Dignified Treatment: intervention required in 20 facilities, with good practice observed in 12 facilities.

- Complaints and Compliments Management: below acceptable standards found in 38 facilities and some good practices in 17 facilities.

Discussion

In discussing the impetus of frontline service delivery monitoring, there is a narrative about the former president having been a listening president. In her seminal paper on this particular subject, Duncan argues that the description of the Jacob Zuma as a listening president is one that is problematic (2010:20):

> "While the Zuma administration is listening to the protestors, there is no evidence yet that they are hearing. Activists are starting to suspect that the newly created points of access to the decision-making system may be designed to deflect oppositional voices rather than to entertain their demands seriously. Responses to this suspicion have varied, leading to a contest for the formal levers of power on the one hand and increasingly militant forms of direct action on the other."

Likewise, as noted in the quote above, the implementation of monitoring systems such as FSDM was the determinant factor on whether citizens would ever realise change.

Despite challenges, notable progress has been made in the implementation of FSDM. According to an early FSDM report, senior management and leadership have shown a discernible improvement in their focus on the frontline, as evidenced by departmental Strategic Plans, Annual Performance Plans, and Budget speeches reflecting this shift. This indicates a growing maturity within national and provincial line departments, recognising the strategic importance of functional frontline facilities (DPME 2015).

The FSDM 2014 guidelines highlighted a tendency to revert to the previous state when monitoring efforts stalled. During FSDM visits to facilities selected for improvements monitoring, it was observed that implemented improvements were not adequately maintained or monitored, resulting in a regression to the baseline state. This lapse is attributed to the absence of frontline monitoring and a lack of accountability for the quality of service provided at the facility level. Additionally, there is a significant likelihood that the findings from FSDM monitoring visits may not be acted upon promptly and decisively to address identified obstacles.

To address these issues, the FSDM initiative is gradually shifting its focus towards monitoring the sustainability of improvements. This entails facilitating and monitoring the implementation of agreed-upon actions. By continuously monitoring facilities selected for improvements, the FSDM programme aims to ensure the long-term viability of the improvements made (DPME 2014).

Table 10.2:   *Monitoring Successes and Challenges*

Monitoring system	Successes	Challenges
FDSM	FSDM has been institutionalised. Several sectors (Justice, SAPS, SASSA, Home Affairs) have developed their own frontline monitoring programmes that monitor service delivery across all facilities. Frontline monitoring has been able to demonstrate the value of collecting views of service users to measure the quality-of-service delivery. Value of on-site verification of the reported results. Culture of M&E is being instilled within facility management as we continue to engage with facilities on implementation of the improvements within their facilities.	Monitoring and evaluation systems in South Africa still evolving – limited capacity in data collection, analysis, dissemination and knowledge management. Inadequate investment on service delivery improvement initiatives (influenced by budget constraints, delegation powers, and compliance monitoring instead of decision-making). DPME is not able to resolve challenges faced by sectors. No consequence management for non-compliance. FSDM is implemented as a stand-alone programme and not adaptive to the changing environment and focuses on real issues.
CBM	Innovative use of existing community programmes / structures (CDWs, ward committees, traditional councils, community work programmes, CPFs, clinic committees etc.). Successfully leveraged on existing local and community structures and programmes, using participatory approaches. This has resulted in strong local ownership and effective use of resources. Identification of root causes through multi-stakeholder dialogues. Bringing government officials, community leaders, and members together to subject problems to a root-cause analysis has provided valuable insights into the underlying causes of service delivery problems. Developing solutions jointly through facilitated dialogue has built a shared purpose between government officials and community members.	Institutionalising the participatory problem-solving approach that underpins CBM work, which is foreign to the way government officials are accustomed to work, and there is currently limited capacity to facilitate this kind of work. Implementing and monitoring commitments made through CBM processes. Internal capacity constraints have prevented the development of effective systems for monitoring commitments made through CBM processes. The CBM initiative has struggled to effectively integrate with other programmes in DPME, to leverage on relationships and capacities.

In her study of the Gauteng Provincial Department of Home Affairs' roll out of FDSM, Mmako (2018) observed that there are gaps in actioning of FDSM findings. She thus noted:

> "Participant responses suggest that there is no standardised approach in actioning the FSDM findings once they have been received from the FSDM team. There is an indication from the findings that most of the FSDM findings are dealt with at the level of the facilities, and only findings that are beyond the scope of facilities are then communicated to the next level managers; that loop continues until the findings reach the district managers, provincial manager and then the DHA national office" (Mmako 2018:106).

In Gauteng DHA, utilisation is employed by political leadership. The FSDM is a tool that has been used by the political leadership in the department to conduct their facilities monitoring. This is a significant opportunity to ensure that FSDM findings are catalysed, as political leadership has the level of authority to drive service delivery blockages (Mmako 2018:104).

Conclusion

The chapter highlights that FSDM played a crucial role as the government's initial step in demonstrating its commitment to realising effective services for citizens. The catalyst for monitoring frontline services was the ongoing civil unrest sparked by poor service delivery during former president Zuma's tenure. A significant policy obstacle was the lack of robust accountability mechanisms and insufficient incentives for accountability. In response to these challenges, the Zuma administration introduced an outcomes-focused approach to government service delivery, necessitating the introduction of various strategies. The DPME, established within the Presidency, took the lead in overseeing these reforms and conceptualising government-wide M&E approaches. Among these approaches, frontline service monitoring was introduced.

While the introduction of FSDM strategically provided public officials with evidence to enhance the quality of frontline services, its capacity to ensure the implementation of its findings remained limited. As the chapter demonstrates, participating public offices were able to improve citizens' service experiences by monitoring and enhancing simple aspects such as visible signage and reduced waiting times. Despite progress made in areas where FSDM was piloted, it is evident that systemic challenges persist at the national level, significantly constraining the influence of such approaches. An illustrative example is the issuance and renewal of driver's licences, which have become sources of frustration for citizens in recent years.

Reference List

Centre for Public Impact. 2016. South Africa's Department of Performance Monitoring and Evaluation (DPME). https://www.centreforpublicimpact.org/case-study/performance-monitoring-in-south-africa

Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K., Rogers, F.H. 2006. Missing in action: teacher and health worker absence in developing countries. *Journal of Economic Perspectives*; 20: pp. 91–116. https://doi.org/10.1257/089533006776526058

Cheruiyot, K., Wray, C., Katumba, S. 2015. Spatial Statistical Analysis of Dissatisfaction with the Performance of Local Government in the Gauteng City-Region, South Africa. *South African Journal of Geomatics*, Vol. 4, No. 3, August 2015. https://doi.org/10.4314/sajg.v4i3.4

DPME. 2014. FSDM Guideline 4.2.7.

DPME. 2014. FSDM Guideline 4.2.8.

DPME. 2015. FSDM Annual Findings Report 2014-2015.

DPME. 2018. Frontline Monitoring and Support Operational Guidelines Framework 2018-2019.

Duncan, J. 2010. Voice, political mobilisation and repression under Jacob Zuma: *A Decade of Dissent: Reflections on Popular Resistance in South Africa, 2000- 319 2010*, symposium, University of Johannesburg, School of Tourism and Hospitality, 12-14 November 2010.

Fraser-Moleketi, G., 2006. *Public Service Reform in South Africa: An Overview of Selected Case Studies from 1994-2004*, Pretoria: University of Pretoria.

Friedman, J.Z. 2014. Sticking to the Numbers: Performance Monitoring in South Africa, 2009-2011.

Fu, H. & Sanchez-Paramo, C. 2020. High-frequency monitoring of COVID-19 impacts. https://blogs.worldbank.org/opendata/high-frequency-monitoring-covid-19-impacts

Goldman, I., Engela, R., Akhalwaya, I.Y., Gasa, N., Léon, B., Mohamed, H., & Phillips, S. 2012. Establishing a national M&E system in South Africa. *PREM Notes; No. 21. World Bank, Washington, DC.* https://openknowledge.worldbank.org/handle/10986/17084

Graham, V. 2015. No longer 'proudly South African'? From 'Rainbow Nation' to 'Dissatisfied Customers': South Africa – 20 years on. Paper prepared for the *British Political Science Association Conference* Sheffield, 30 March–1 April 2015. https://www.psa.ac.uk/sites/default/files/conference/papers/2015/PSAVGrahampaper.pdf

Gumede, W. 2017. Citizen, civil society and media-based monitoring and evaluation. Policy Brief 19 https://democracyworks.org.za/policy-brief-19-citizen-civil-society-and-media-based-monitoring-and-evaluation/

Hood, C., 1991. A Public Management for all Seasons?. *Public Administration,* 69(1), pp. 3-19. https://doi.org/10.1111/j.1467-9299.1991.tb00779.x

International Institute for Environment and Development (IIED). 2019. South Africa's national development monitoring and evaluation. https://www.iied.org/south-africas-national-development-monitoring-evaluation

Kettner, P. M., Moroney, R. M. & Martin, L. L. 2013. *Designing and Managing Programs: An Effectiveness-Based Approach*, 4th edn, Thousand Oaks, CA, Sage.

Kuye, J. & Ajam, T., 2012. The South African Developmental State Debate Leadership, Governance and a Dialogue in Public Sector Finance. *African Journal of Public Affairs,* 5(2), pp. 48-67.

Leach, L.M. 2018. *A framework for factoring citizen perceptions in local public service value systems to improve service quality*. PhD Dissertation, Faculty of Economic and Management Sciences, Stellenbosch University.

Mmako, D. 2018. *Understanding officials' perspectives on the effectiveness of Frontline Service Delivery Monitoring in improving service delivery in the Gauteng Department of Home Affairs.* Master of Management Research Report, University of the Witwatersrand.

Muralidharan, K. Das, J., Holla, A., and Mohpal, A. 2017. The fiscal cost of weak governance: Evidence from teacher absence in India, *Journal of Public Economics*, Volume 145, 2017, pp. 116-135. https://doi.org/10.1016/j.jpubeco.2016.11.005.

National Planning Commission (NPC). 2012. *National Development Plan 2030: Our Future - Make it Work,* Pretoria: Presidency of South Africa.

Osborne, D. & Gaebler, T., 1992. *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector.* Reading: Addison-Wesley Publishing Company.

Plangemann, K. 2016. Strengthening performance monitoring and evaluation, in Alam, Asad, Renosi Mokate, and Kathrin A. Plangemann, editors. 2016. *Making It Happen: Selected Case Studies of Institutional Reforms in South Africa.* Directions in Development. Washington, DC: World Bank. doi:10.1596/978-1-4648-0768-8.

Porter, S., & Goldman, I. 2013. A Growing Demand for Monitoring and Evaluation in Africa. *African Evaluation Journal*, 1(1), 9 pages. doi:https://doi.org/10.4102/aej.v1i1.25

Power, M., 1997. *The Audit Society: Rituals of Verification.* https://www.demos.co.uk/files/theauditexplosion.pdf

Republic of South Africa. 1997. *The White Paper on the Transformation of the Public Service. Government Gazette, vol 388, no. 18340. General Notice 1459 of 1997. Department of Public Service and Administration,* Pretoria: Government Printer.

Phillips, S., Goldman, I., Gasa, N., Akhalwaya, I. & Leon, B. 2014. A focus on M&E of results: an example from the Presidency, South Africa, *Journal of Development Effectiveness*, 6:4, pp. 392-406, DOI: 10.1080/19439342.2014.966453

The Presidency. 2010. President Zuma to report back to residents of Siyathemba Township. https://www.thepresidency.gov.za/content/president-zuma-report-back-residents-siyathemba-township

UNECA. 2010. *Innovations and Best Practices in Public Sector Reforms: The Case of the Civil Service in Ghana, Kenya, Nigeria and South Africa,* Addis Ababa: UNECA.