BUILDING SOCIAL PROTECTION FLOORS IN AFRICA

TRANSFORM

MONITORING & EVALUATION

TRANSFORM is the result of an iterative process of co-creation involving experts and practitioners from southern and eastern Africa. This manual is based on a document prepared by Andrew Kardan, Valentina Barca, and Andrew Wyatt (Oxford Policy Management) with contributions from Franck Earl (SASSA), Luis Frota, and Thea Westphal (ILO). The revised content on gender-responsive and disability-inclusive administration of social protection was prepared by Tia Palermo, Maja Gavrilovic, Morgon Banks, and Lusajo Kajula from Policy Research Solutions (PRESTO) LLC and reviewed by the TRANSFORM Advisory Group.

The editors of the TRANSFORM curriculum series are Luca Pellerano, Luis Frota, and Nuno Cunha. Participants to workshops in Kenya, Zambia, and Tanzania provided useful comments and inputs. The content of this manual does not reflect the official position of the different organizations supporting the TRANSFORM initiative.

For further information, you can contact the TRANSFORM initiative at

transform_socialprotection@ilo.Org or visit **www.transformsp.org**

Transform, (2023) "Governance of Social Protection Systems: Institutions, Organizational Structure and Capacity - Manual for a Leadership and Transformation Curriculum On Building and Managing Social Protection Floors in Africa", (Original publication date: 2017; Updated: 2023) available at TRANSFORM | Building Social Protection Floors in Africa **www.transformsp.org**

Suggested Citation: Transform, (2023) "Governance of Social Protection Systems: Institutions, Organizational Structure and Capacity - Manual for a Leadership and Transformation Curriculum On Building and Managing Social Protection Floors in Africa", (Original publication date: 2017; Updated: 2023) available at TRANSFORM | Building Social Protection Floors in Africa (**www.transformsp.org)**

CONTENTS

LIST OF ABBREVIATIONS

BFP	Bolsa Familia Programme
CBM	Citizen-Based Monitoring
CLIC	Community LEAP Implementation Committee
CONEVAL	National Council for the Evaluation of Social Development Policies
CRF	Core Result Framework
CSG	Child Support Grant
CSOs	Civil Society Organisations
DPME	Performance, Monitoring and Evaluation in the Presidency
DSW	District Social Welfare
DSWO	District Social Welfare Officer
FSP	Food Safety Pack
GEWEL	Girls Education Women's Empowerment and Livelihood
GRZ	Government of the Republic of Zambia
HIV	Human Immunodeficiency Virus
IGD	Decentralized Management Index
ILO	International Labour Organisation
IMCs	Independent Monitoring Checks
INAS	National Institute for Social Action
LEAP	Livelihood Empowerment Against Poverty
LFA	Logical Framework Approach
LMU	Leap Management Unit
M&E	Monitoring and Evaluation
MCI	Independent Community Monitoring
MDS	Ministry of Social Development
MIS	Management Information Systems
NGOs	Non-Governmental Organizations
NSPP	National Social Protection Policy
NSPS	National Social Protection Strategy
OECD	Organisation for Economic Cooperation and Development
OMF	Operational Management Framework
PM	Performance Management
PSCM-PS	Civil Society Platform for Social Protection
PSM	Propensity Score Matching
PSNP	Productive Safety Nets Programme
PSSB	Basic Social Subsidy Programme
PWDs	People Living with a disability
R 202	ILO Social Protection Floors Recommendation, 2012 (No. 202)
RCT	Randomized Controlled Trials
RDD	Regression Discontinuity Design
RSWO	Regional Social Welfare Officer
SASSA	South African Social Security Agency
SCT	Social Cash Transfer
UNCRPD	United Nations Convention on the Rights of Persons with Disabilities
UNICEF	United Nations Children's Fund

1

MONITORING AND EVALUATION (M&E) SYSTEMS

1.1 WHY M&E IS IMPORTANT FOR EFFECTIVE SOCIAL PROTECTION PROGRAMMING

The ILO's R 202 recommends that **countries should regularly 'collect, compile, analyze, and publish an appropriate range of data statistics and indicators'**.[1] This is critical to safeguarding compliance with existing legislation, ensuring transparency and accountability and building a basis for the continuous improvement of social protection systems and programmes to meet the specific needs of and the equitable inclusion of vulnerable groups, including women and girls, and people with disability.

A good M&E system promotes a continuous learning cycle, fosters transformation in social protection, and improves service delivery. Ideally, it is triggered by a continuous demand for M&E and gives equal importance to monitoring and evaluation functions (see Box below). Moreover, an M&E framework that harmonizes indicators from across social protection programmes can help to overcome potential fragmentation at the policy and programme level, while reaping benefits in terms of cost and capacity synergies.[2]

Specifically, a well-functioning M&E system in the social protection sector can[3]:

1. Improve policy/programme management and planning ('inwards facing' M&E)

- Improve policy/programme design: to learn about the efficiency and effectiveness of a policy/programme for decisions on whether to extend, improve, or eliminate it. The ultimate aim would be to better serve the poor and vulnerable and provide more equitable and efficient services.
- Help solve problems in policy/programme implementation: monitoring execution to detect and correct implementation problems and facilitate evidence-based fine-tuning of the operational design.
- Help prioritize, plan and budget: helping relevant authorities and managers to coordinate and prioritize activities and undertake planning and budget allocation decisions.

1 ILO, 2012, Recommendation 202, Section II
2 Government of Kenya, p. 83
3 Reorganised by Attah et al (2015)

2. Enhance policy/programme accountability ('outwards facing' M&E)

- Ensure accountability within the government: monitoring of policy/programme execution to ensure that agents are doing what they have undertaken to do.
- Provide public information for external accountability: providing information to elected officials and the general public to

i. legitimize the policy/programme through the provision of results and achievements;

ii. encourage public choice and voice.

- Track outcomes using an equity lens, disaggregating data by age, gender, disability status, ethnicity, and other known vulnerabilities to ensure the programme is equitably reaching all recipients and leading to positive impacts.

When discussing M&E systems for Social Protection, the standard approach is to lump the two concepts of monitoring and evaluation together, without necessarily distinguishing between the very different objectives these two activities help to achieve.

According to the Organisation for Economic Cooperation and Development (OECD, 2002) and to the commonly accepted DAC terminology, monitoring can be defined as a "continuous function that uses the systematic collection of data on specified indicators to provide management and the main stakeholders of an on-going development intervention with indications of the extent of progress and achievement of objectives, and progress in the use of allocated funds".

Evaluation, on the other hand, is defined as the "systematic and objective assessment of an on-going or completed activity, programme or policy, its design, implementation, and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact, and sustainability"

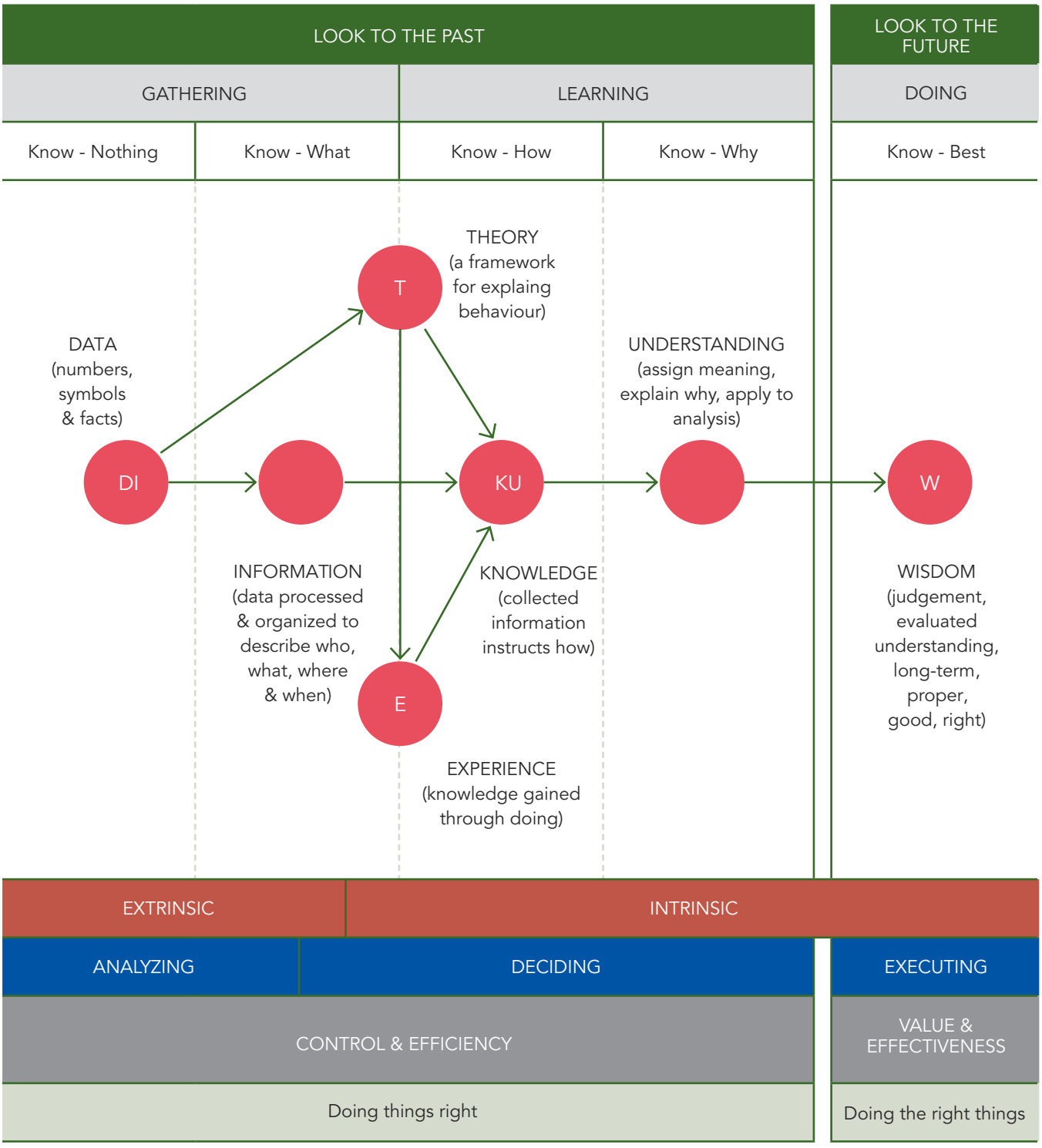
- Monitoring focuses primarily on the relationship between inputs and outputs, with a view at "improving" the efficiency of the implementation.
- Evaluation focusses primarily on the relationship between outputs and impacts, with a view at "proving" the effectiveness of the design.

M&E data is not of use in itself unless it is translated into information and knowledge and ultimately shapes decisions (see Figure 1). Developing an M&E system is about building capacity and practice to gather information from the past course of action, learn from experience, and use data to orient and improve the course of action in the future.

The role of M&E is to provide reliable information to enable the decision makers "do the right things" as well as to "do things right". An M&E system should therefore provide information that allows it to:

- **improve control and efficiency** of social protection systems (relating primarily to the "internal" objectives of M&E mentioned above)
- **prove value and effectiveness** (relating primarily to the "external" objectives of M&E mentioned above).
  - **Avoid bias and track progress on inclusive and context-specific indicators** for vulnerable groups
  - **Detect unintended effects** of programmes on recipients and identify gaps and weaknesses in programme delivery to modify the design and/or implementation practices to avoid harmful outcomes and improve results

Figure 1: From Data to Decisions



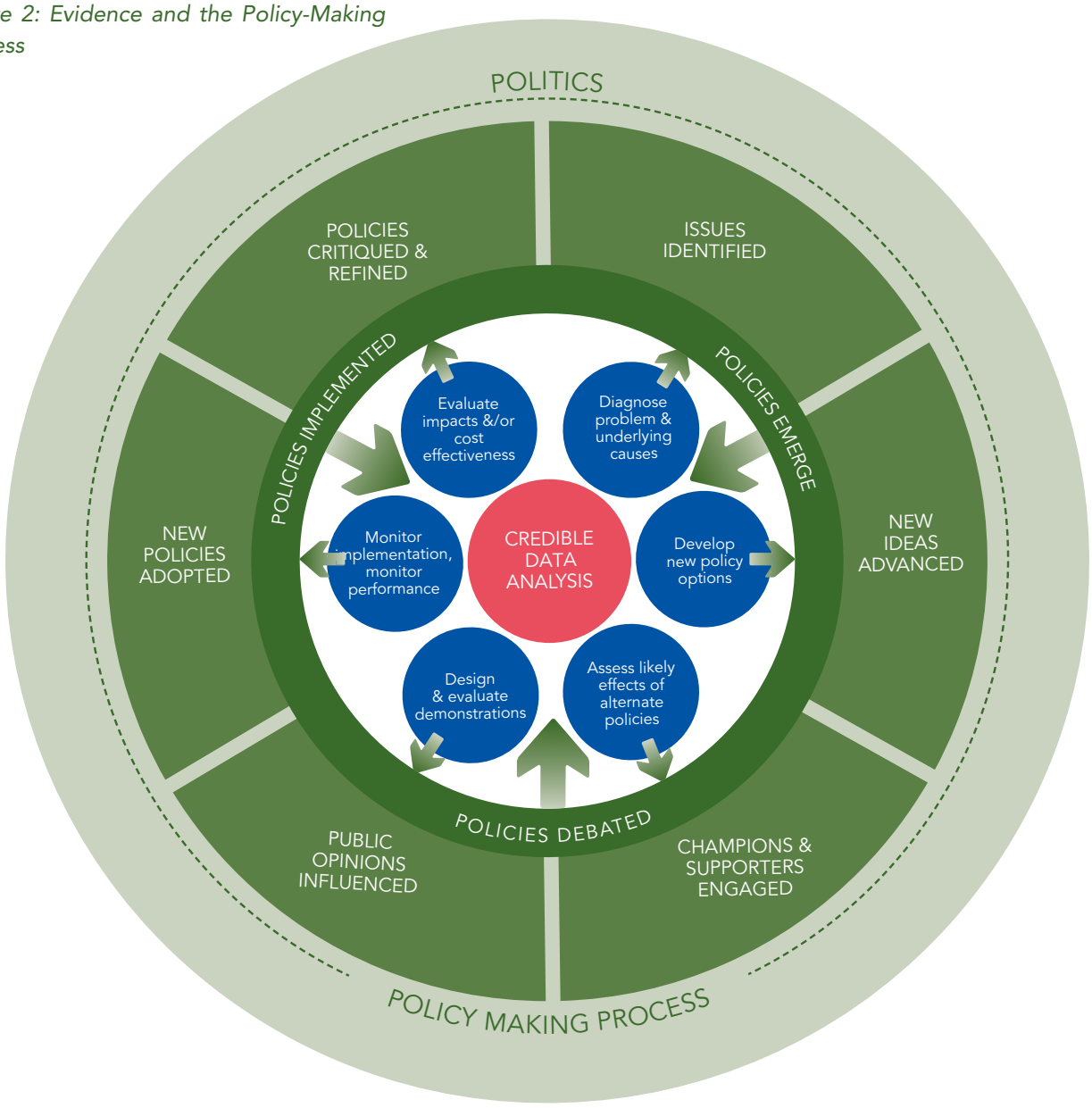
Source: adapted from Ackoff (1989)

1.2 POLITICAL ECONOMY OF MONITORING AND EVALUATION

1.2.1 M&E and the Policy-Making Process

The growing focus on "evidence-based policy making" reflects the increasingly central role of credible data and analysis at all steps of the policy-making progress. Monitoring and evaluation instruments can play different roles at different stages of the policy design and implementation process (see Figure 2).

Figure 2: Evidence and the Policy-Making Process



Source: Urban Institute - http://urbn.is/2gQc3hU

Yet, the core business of policy makers is to make decisions about whether credible evidence is available or not. The learning process is shaped by evidence as much as it is shaped by theoretical assumptions and suppositions, as well as previous experience (refer to steps T and E in Figure 1). Moreover, the evidence provided by M&E systems is in direct competition with layman's opinions, gossip, hearsay, anecdotes, long-held prejudices and beliefs. Even in the presence of unbiased high quality M&E information, decisions can be made based on prior opinions, perceptions and experience, in addition to, or even disregarding the evidence available.

Decision-making is a complex process which brings together a constellation of many variables including political considerations. Decisions are not made in a linear and necessarily rational manner and usually are a product of several interests, influences, ideas, and agendas working in isolation or tandem. To a large extent, **the success or failure of an M&E system depends on the interplay between evidence and other factors in shaping the policy-making and policy-implementation process.**

Political considerations intrude in three major ways, and the evaluator who fails to recognize their presence is in for a series of shocks and frustrations (Weiss, 1970):

- First, the policies and programmes with which M&E deals are the **creatures of political decisions.** They were proposed, defined, debated, enacted, and funded through political processes, and in implementation, they remain subject to pressures-both supportive and hostile-that arise out of the play of politics
- Second, because M&E is undertaken to feed into decision-making, its reports enter the political arena. The evaluative evidence of programme outcomes has to compete for attention with other factors that carry weight in the political process
- Third, and perhaps least recognized, **M&E itself has a political stance.** By its very nature, it makes implicit
- political statements about such issues as the problematic nature of some programmes and the unchallenged ability of others, the legitimacy of programme goals and programme strategies, the populations of interest (through data disaggregation), the utility of strategies of incremental reform, and even the appropriate role of the social scientist in policy and programme formation.

Box 1: The Role of M&E in influencing the SCT scale-up in Zambia

In October 2013, the Zambian Parliament discussed and approved the 2014 government Social Cash Transfer Budget, which represented a sheer eightfold increase in the government budget to the programme— ZMW 150 million, up from ZMW 17.5 million in 2013.

In retrospect, it is evident that a number of factors came together to drive the scale-up decision. No single factor was by itself sufficient to trigger the increase. It was the coming together of a number of factors in time that created an environment conducive for an increased budget allocation.

First, the new Patriotic Front (PF) government that had been elected in November 2011 put strong emphasis on social protection approaches to tackle poverty and inequality. Its 2011–16 Manifesto has a dedicated chapter on Social Protection, which announces an NSPP and mentions the importance of increased government budget allocations for sustainability of the cash transfer programme.

Secondly, the sustained criticism of Zambia's subsidy programmes for fertilizer, maize, and fuel created an opportunity to channel resources to pro poor programmes as pointed out by the Presidents Office.

Thirdly, the Ministry started the drafting process of the policy. The Ministry also made a presentation to the Secretary to the Cabinet and a number of permanent secretaries in May 2013 and covered the findings from the Child Grant impact evaluation.

Fourthly, the Secretary to the Cabinet played a championing role and confirms the often advanced thesis that major policy shifts require a political 'champion' Some analysts have attributed the 'triggering' of the scale-up decision to the favourable attitude of the President and his willingness to 'push' the social protection agenda.

Impact Evaluations significantly contributed to the enhancement and reputation of the cash transfer among key audiences including the Ministry of Finance. Secondly the evaluation findings helped to deepen understanding of the key audiences and had them begin to shift their minds from transfers being a consumption expenditure item to it being an investment.

Source: Authors based on Davis et al (2016)

Box 2: Political Pressure unlocks resources for Social Protection

Upon taking power in 1994, Nelson Mandela's new government set out to rebuild a nation that had been torn apart by apartheid. Among its first moves, the new government looked to address widespread poverty and reform unfairly distributed social services. Reforming the child welfare system became a top priority, made even more urgent by the growing burden of AIDS within already disadvantaged communities.

To kick-start reform, the minister of welfare and provincial leaders convened experts through the Lund Committee on Child and Family Support. In 1997, the committee proposed a new social welfare programme, the Child Support Grant (CSG), that would provide ZAR70 (US$15) each month to the poorest 30 percent of children under seven years old. The South African Parliament approved the committee's recommendations with one notable modification: an increase in the grant's initial value to ZAR100 (US$21). Conditions for receipt of the grant included participation in development programmes and proof of immunization status.

Political pressures pushed the government to expand this programme incrementally. 'Affordability' was a constraint at times, but not at others. The government's commitment to the CSG, bolstered by civil society organizations that actively support the programme, has been vital to its expanding coverage and impact. When the programme was first launched, the finance minister, Trevor Manuel, worried that the system was unsustainable and would turn South Africa into a welfare state. Over time he became a firm supporter. The evidence clearly shows its benefits for human development and confirms its critical role within the government's broader strategy to roll back structural inequality. The level of public investment in the social grants overall, and the CSG in particular, is a testament to the government's dedication.

The Child Support Grant was adopted and implemented in a short space of time in a contracting fiscal milieu amid skepticism from civil society mainly due the Politicians who were able to garner support from their constituents about the need to promote social and economic justice and to address the legacy of its apartheid past.

Source: Patel (2011)

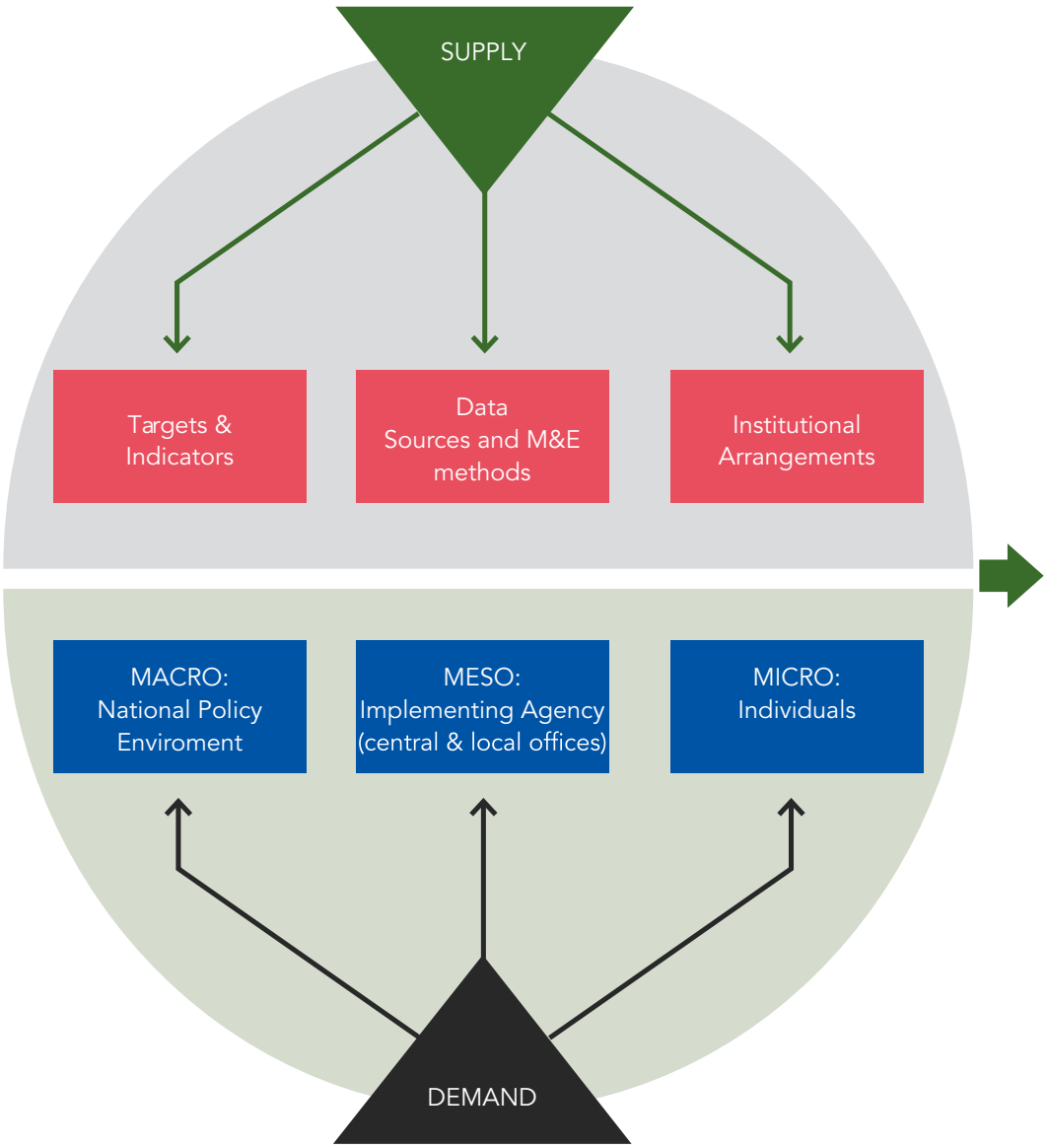
1.3 M&E EVIDENCE: SUPPLY AND DEMAND[4]

To achieve the objectives outlined above, M&E systems must be designed in such a way to strike a balance between the supply (the capacity needed to produce quality evidence in a timely fashion, and the cost of doing so), and the demand for different types of evidence needed by multiple users for decision making. We discuss how this can be ensured in the remainder of this document.

The supply of good information must be matched by effective demand and the use of evidence. These two "forces" mutually enforce one another, resulting in better design of M&E systems for social protection. Data users must know what evidence they need and why, whilst the data providers must know how to generate and disseminate quality information. Supply and demand of M&E information must work in sync to avoid a mismatch. M&E information must be provided equitably and not discriminate against stakeholders based on their gender, age, disability status, and other vulnerabilities. This includes making results available and disseminating them in simple, user-friendly formats, and accessible formats such as Braille, audio, and video with sign language.

4 This section draws largely from Attah et al (2015) and Segone (2008)

Figure 3: Demand and supply of M&E data for social protection



Source: Attah et al (2015)

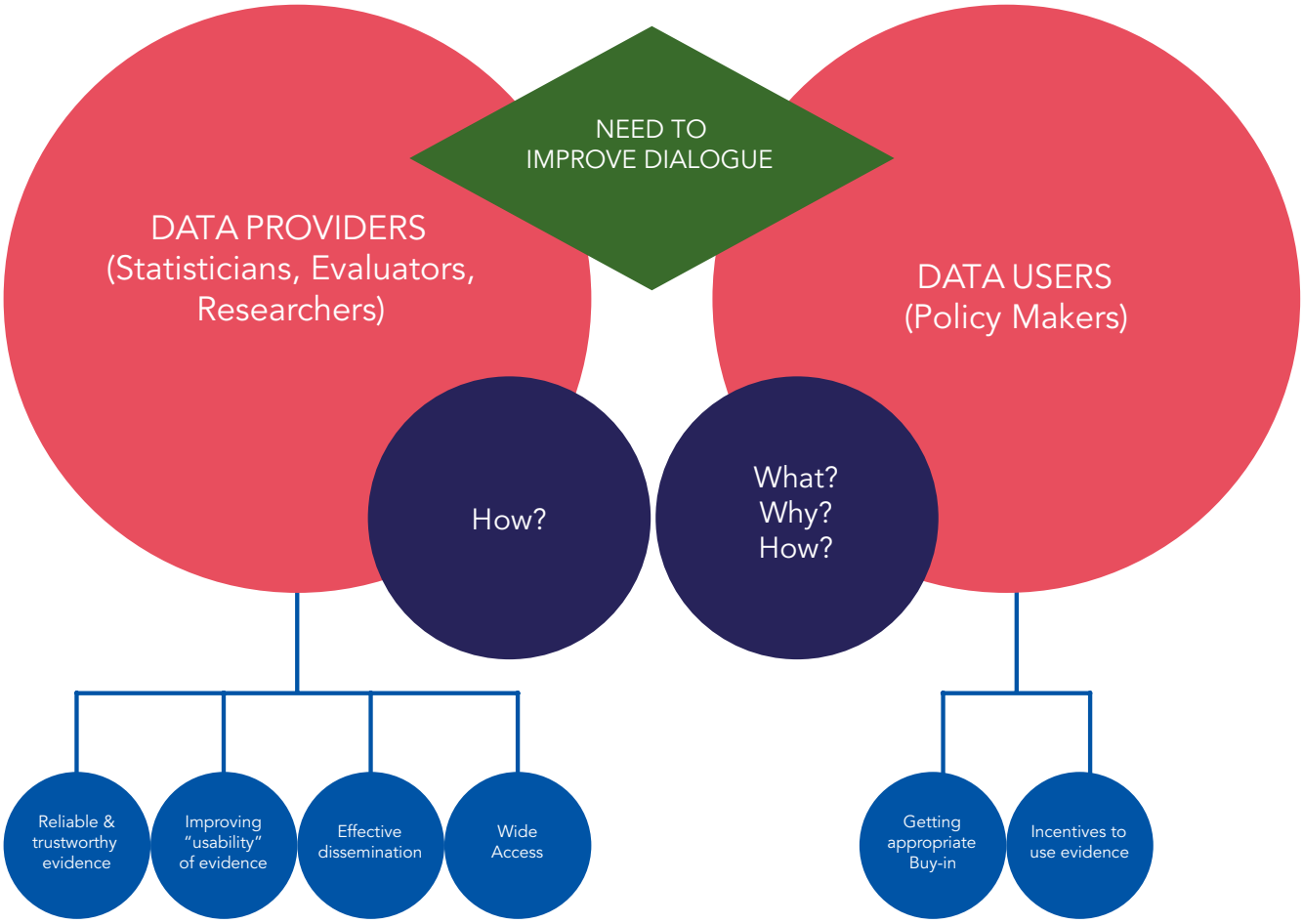
If evidence that is technically sound is not policy relevant, then it will not be used by policymakers. The opposite also applies, that is, policymakers may be forced to use poor quality evidence if this is the only evidence available that addresses their policy questions. Getting the right balance between both the principles of professional autonomy and accountability and the relevance of the evidence produced, is paramount. (Segone, 2008)

The following key factors are critical to ensure an effective balance between the demand and supply of M&E. They are further discussed in Section 3.

- **Improve the dialogue between policymakers and evidence providers.** To maintain and sustain this balance, deliberate efforts must be made to ensure that there is a constant dialogue between the data providers and data users. This is strategic because, at the end of the day, policymakers know what evidence they need, why they need it, and when they need it. Statisticians, evaluators, and researchers know how to provide that evidence. (Segone,2008)
- **Making evidence "usable" for the policy-making community.** Getting the policy makers to own the evidence needed for the effective implementation of policies is critical. Evidence should not be the property of the data gatherers. The evidence supplied must be reliable, trustworthy, disaggregated to represent the situation of vulnerable groups, and well disseminated in user-friendly and accessible formats with wide access for various users and interest groups. A key issue is how to communicate the findings to those who need to know.

- **Building and supporting the national and local capacity to gather and analyse data.** This promotes local and national ownership of the evidence and enhances political buy-in from policymakers. The source of data matters to policymakers.
- **Provide incentives to use evidence.** A key ingredient for ensuring that policy makers are using the evidence is to create an incentive structure that will increase the uptake of evidence. This comprises different strategies (see more on carrots, sticks and sermons in section 3), all hinging around the notion of promoting good performance based on demonstrable (evidence-based) results. Partnerships with civil society representatives of different stakeholder groups may increase the quality of the data collected and the likelihood of evidence being used by decision-makers, as civil society organisations can share the evidence through their networks and are often effective at lobbying for change. They may also act as gatekeepers to ensure that the rights of vulnerable groups – such as women and girls and people with disability – are included.

Figure 4: Increasing use of evidence by balancing demand and supply



Source: Segone (2008)

Many governments and organizations are moving from "opinion-based policy" towards "evidence-based policy", and are in the stage of "evidence-influenced policy". This is mainly due to the nature of the policy environment as well as the national technical capacity to provide good quality and trustworthy evidence. Box 3 below describes four possible situations.

Box 3: From Opinion Based to Evidence Based policy making

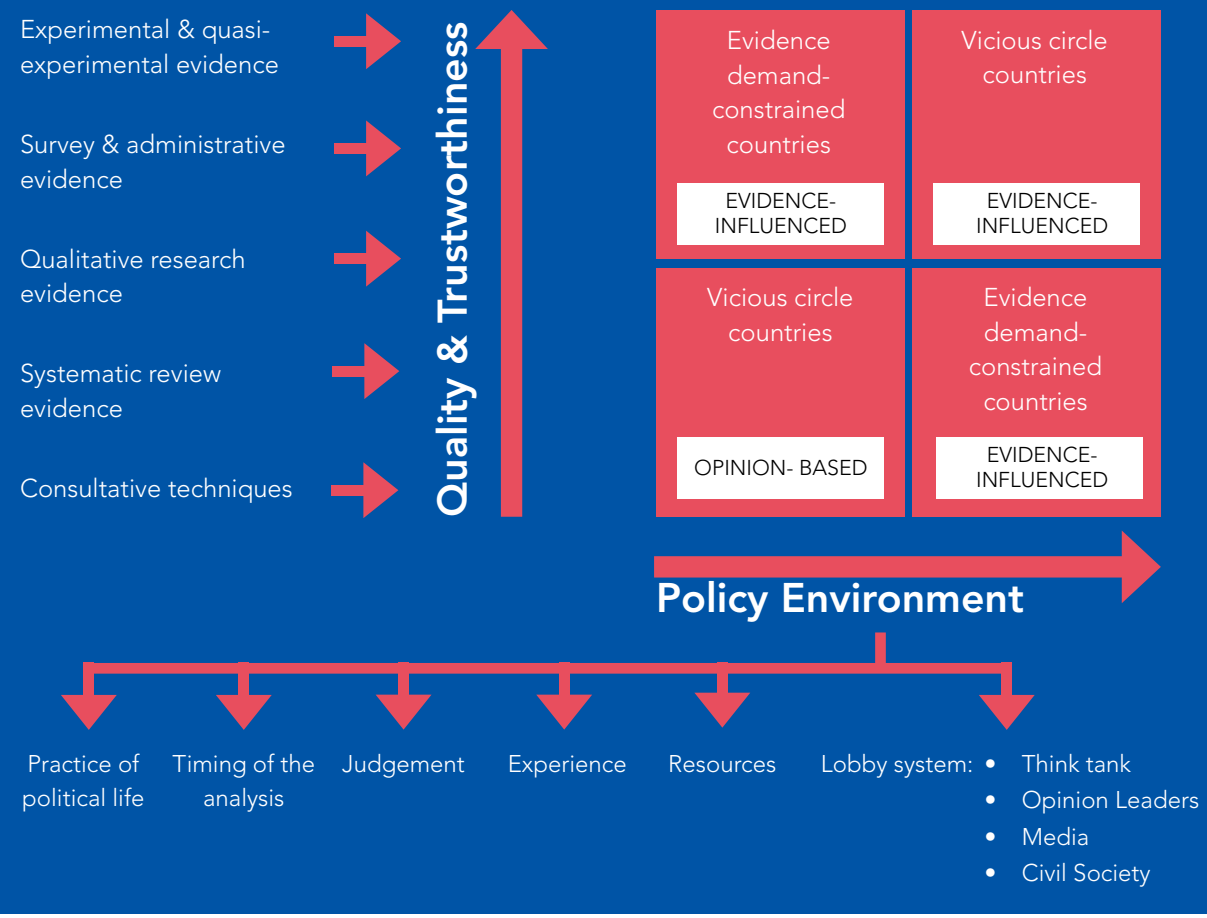
Vicious circle countries. Evidence is weak and policymakers make little use of it. Evidence-based policy-making is not practised, which results in poor policy decisions and poor development outcomes. In this case, it is necessary to adopt measures which will simultaneously increase both the demand and supply of evidence, as well as improve the dialogue between producers and users of evidence.

Evidence supply-constrained countries. Although evidence is weak, it is increasingly used by policymakers. However, evidence deficiency reduces the quality of decision-making which results in poor development outcomes. Policymakers are likely to resent being held to account based on inadequate evidence. The priority is to adopt measures to increase the quantity and quality of evidence, which will require additional technical assistance for capacity development, as well as to improve the dialogue between producers and users of data. The challenge is to strike a balance between generating improvements to evidence quickly while laying the foundations for better performance of the national monitoring and evaluation system in the long run.

Evidence demand-constrained countries. The quantity and quality of evidence are improving, but it is not used for decision-making because policymakers lack the incentives and/or the capacity to utilize it. In this case, priority should be given to the adoption of measures to increase the demand for evidence, as well as to improve the dialogue between producers and users of data.

Virtuous circle countries. The production of good (or at least improved) evidence is matched by its widespread (or at least increased) use in decision-making. These two processes mutually reinforce each other, resulting in better policy design and better development outcomes.

Figure 5: Dynamic of Policy making



Source: Segone (2008)

1.4 UNDERSTANDING INFORMATION NEEDS

Critical to the successful uptake of evidence generated through M&E is to assess the stakeholders' various interests, specific information needs and the influence that they wield and the incentives at play.

Three major categories of stakeholders can be distinguished from social protection systems[5]:

- the **national authorities** with its various components (the executive, the legislative, the control, and oversight bodies), both at national and at decentralized government levels. These different national stakeholders may have discordant interests and cannot be considered as one homogeneous group;
- national **civil society** (which is also not a homogeneous group and can include NGOs, churches, research institutes, women's groups, Organisations of Persons with Disabilities, programme recipients as well as the wider public) recipient. They need to have access to programme-related information and evidence, as they can act as a key source of support for social protection policies and their equitable expansion.
- the **international community** (donors and cooperating partners), where relevant. This is also a diverse group and interests do not always converge.

In terms of M&E accountability and learning each of these groups has its interests (see Table 1 below).

Table 1: Key stakeholders - standard information needs

LEVEL	STAKEHOLDERS	INFORMATION NEEDS
Central	Members of Parliament	Mainly interested in information on their constituency. Want to know about the impact and scale-up plans.
	Ministry of Finance	Interested mostly in budget and efficiency/effectiveness of the programme, as well as the impact
	Other Ministries	Interested in resource allocation and impact (especially when related to their core area); some interest in coordinating operations
	Donors	Strong focus on impact, sustainability, Value for Money, efficiency and effectiveness of operations and overall accountability
	Media	Want to know what is happening when; often information misused for sensational reporting
	Planning Unit within the lead Ministry	Mostly interested in information for planning and budgeting purposes (number of recipients, total amounts disbursed, etc.) as well as ad-hoc responses to parliamentary queries.
	Management Unit within the lead Ministry	Focus on all information above + indicators useful for programme management (cost-efficiency and effectiveness, compliance with Service Standards, etc.)
Province/ district	Provincial/district authority	Interested in impact and overall number and types of recipients at province/ district level; information for coordination and management of lower levels: staff, budget, and Quality Control;
Community	Community-level leaders, social workers; Civil Society Organisations	Interested in the number and identity of recipients in their area and any other information to hold the programme accountable (e.g. citizen perceptions of the programme). Civil society organisations may be issue-specific, for example, those who focus on the rights of people with disability or women and girls.
	Recipients	Recipients often include vulnerable groups such as women, girls, and people with disability. They need to know the transfer value and periodicity.

Source: Authors

5 Bamberger (1991)

The following questions are critical in assessing the relevance of M&E for different stakeholders:

- What decisions, if any, is evidence from the M&E system expected to inform? What would stakeholders do differently because of the evidence provided by the M&E system?
- When would decisions be made? When must M&E information be available to be timely and influential?
- What information is needed as a priority to inform decisions?
- Who will use the evidence from the M&E system, that is, who has the willingness, authority and/or ability to put learning from the M&E system to use?
- Can data be disaggregated by all relevant recipient groups to allow for the monitoring of equity and equality?
- How will the outcomes for different recipient groups – including those who are marginalized and vulnerable – be collected, reported, and disseminated?

1.4 TAKE-AWAY LESSONS

- Recommendation 202 suggests that countries should regularly 'collect, compile, analyze, and publish an appropriate range of data statistics and indicators'.
- Decision-making is a complex process involving different variables including political interests and this process does not happen linearly. As the capacity to generate and use quality evidence increases, many governments are moving from opinion-based policy making to evidence-based policy making.
- Good M&E is critical to monitoring compliance with existing legislation, ensuring transparency and accountability (both internal and external) and building a basis for the continuous improvement of social protection systems (improve policy/programme design; solve problems in policy/programme implementation; help prioritize, plan and budget).
- M&E systems perform two very different functions: they provide evidence for both proving that the programme is "doing the right things", and for improvements to ensure the programme is "doing things right". Both functions should be given adequate importance.
- To achieve its objectives, M&E systems must be designed in such a way to strike a balance between the capacity (and cost) to produce evidence in a timely fashion and with quality – i.e., supply data, and the demand for evidence of a particular kind or nature for decision making according to the needs of multiple users.
- M&E systems need to collect, analyse and report data disaggregated by all relevant recipient groups. Indicators of gender, disability, and age – amongst others – are universally relevant and need to be reported in all settings.
- Reporting needs to highlight gaps in equity and equality for different recipient groups.

2

ENSURING SUPPLY OF M&E DATA[6]

Indicators, M&E approaches, data sources and institutional arrangements, discussed in depth below, are the main building blocks of an M&E framework – helping to track and evaluate programmes against their Theory of Change and related Results Framework.

2.1 DEVELOPING A THEORY OF CHANGE

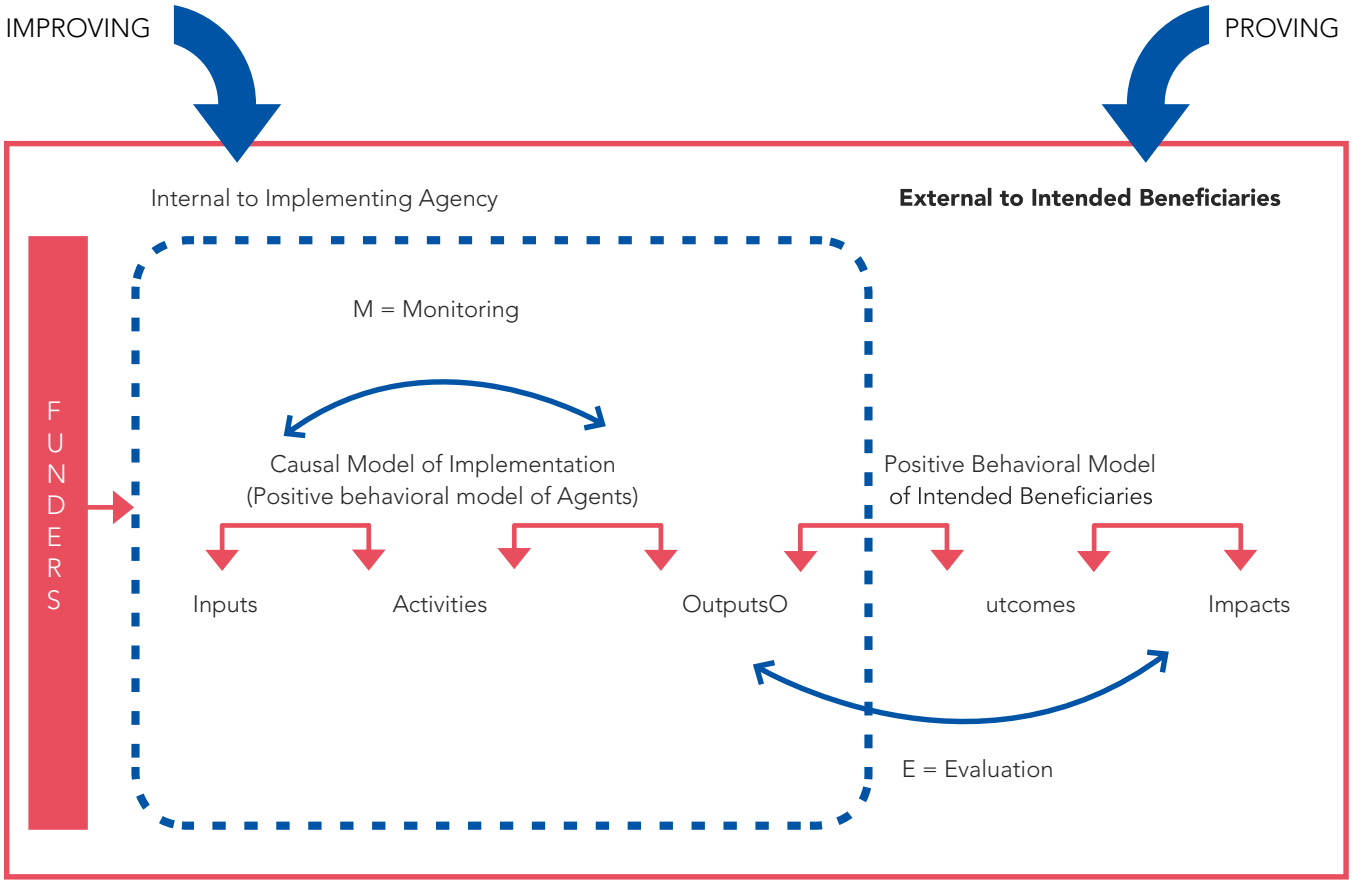
The Logical Framework Approach is an analytical process and set of tools used to support project planning and management. According to the World Bank (2000), "the Logical Framework has the power to communicate the essential elements of a complex project clearly and succinctly throughout the project cycle. It is used to develop the overall design of a project, to improve the project implementation monitoring and to strengthen periodic project evaluation". It provides a set of interlocking concepts which are used as part of an iterative process to aid structured and systematic analysis of a project or programme idea or its **Theory of Change (ToC).**

Standard practice, as articulated in project planning or logical framework approaches, describes a 'development project' as inputs (financial and other resources), which are translated by an implementing agency into specified activities to produce useful outputs. These outputs have the goal of outcomes and impacts of higher well-being for the intended recipients (Pritchett, 2013). See Figure 6 below for a schematic representation of the logical framework structure.

Inputs, activities and outputs are under the control of the programme managers, as their relationship depends on the implementation model, capacity and organisation. Outcomes and impacts are outside the control of the programme managers, as they depend on contextual factors and the behavioural response of intended recipients (and other actors) to the intervention. A robust ToC should be a primary tool for guiding the M&E of interventions (e.g., policies, single and integrated programmes) and assessing changes (positive and unintended) and pathways though which these changes occur and manifest themselves.

6 T his section draws largely from Attah et al (2015)

Figure 6: Logical Framework and the role of M&E



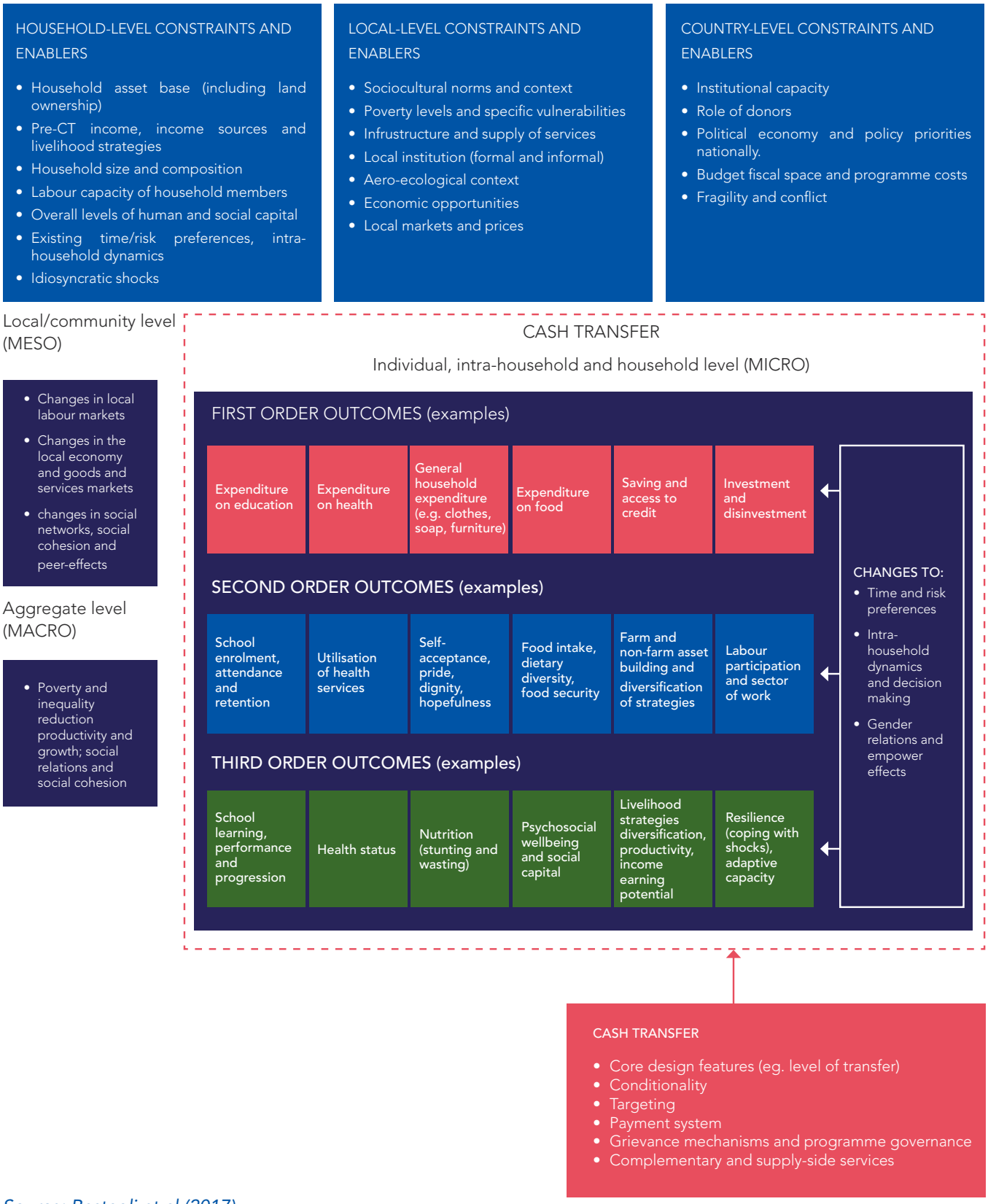
Inputs (what is made available to the project)	Activities (what the project does)	Outputs (achievements that will lead to outcomes)	Outputs (changes external to the project)	Impacts (long run impact on well-being)

A representation of a stylized theory of change for a social protection programme is provided in Figure 7 and a concrete example of a log frame, with a definition of inputs, outputs outcomes and impacts for a cash transfer intervention in Ghana is discussed in Box 4.

The best practice in developing a ToC is to use a participatory approach with key stakeholders, including representatives from vulnerable groups, to ensure important inputs and outcomes are not omitted from M&E data. For example, people with disability face additional challenges accessing social protection compared to people without disabilities, due to physical, financial and attitudinal barriers. These barriers should be considered within a ToC and outcome measures such as programme coverage should be disaggregated by disability status. Similarly, indicators of programme impact should be disaggregated by disability status, and additional disability-specific indicators may be needed to capture areas of concern for people with disability and their households. For example, indicators on the extent to which cash transfers improve access to disability-related services (e.g., assistive devices, rehabilitation and specialist health services, inclusive education) and household dynamics (e.g., time spent on care responsibilities by other household members) should be included.

Similarly, all social protection programmes – even those without explicit gender objectives – are likely to produce gender-specific outcomes. These prospective gender changes should be properly reflected in the programme's ToC (although this rarely systematically happens in practice). This can be achieved by specifying gender aspects in the existing ToC diagrams and narrative (see an example of Ghana's LEAP below) or by reconstructing a specific gender-sensitive ToC from the existing logical frame or programme theory. For example, during the development of a study design to guide a gender-sensitive process evaluation of Ethiopia's Productive Safety Net Programme, the researchers had to reconstruct the existing results framework to explain how various programme activities are expected to produce gender-sensitive changes and the role of programme delivery in mediating these results. Involving people with disability and women in the design of a ToC and the development of M&E indicators can help to (i) identify a more meaningful and accurate vision of change for these groups resulting from an intervention and (ii) map opportunities and challenges to achieving the programme's desired outcomes amongst these groups, as well as appropriate indicators for M&E.

Figure 7: A stylised cash transfer conceptual framework



Source: Bastagli et al (2017)

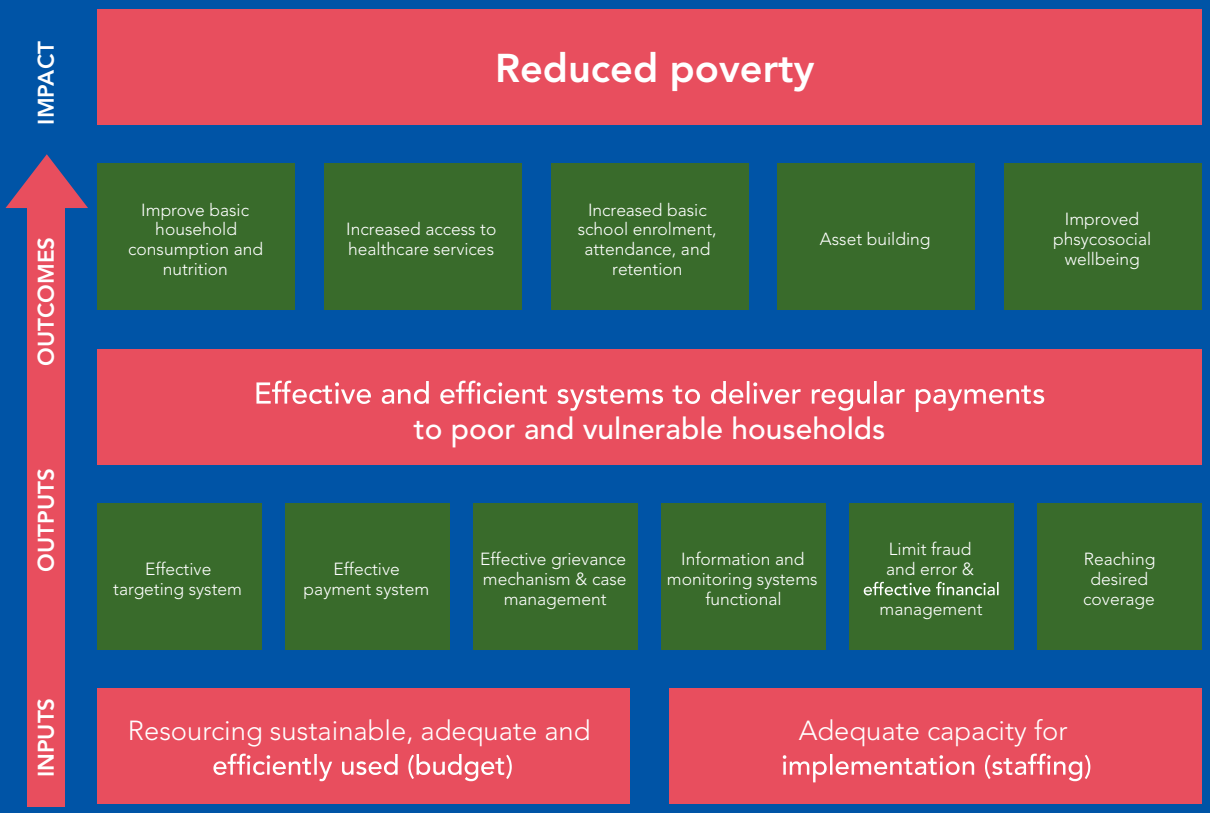
Box 4: Ghana LEAP results framework from inputs to impacts

The Livelihood Empowerment Against Poverty (LEAP) programme is the flagship programme of Ghana's National Social Protection Strategy (NSPS) that began in 2008. The programme was designed to fight poverty among extremely vulnerable populations by providing bimonthly cash payments to extremely poor households with orphans and vulnerable children, the elderly with no productive capacity, and people with severe disabilities with impaired ability to work. The specific objectives of LEAP are:

- To improve basic household consumption and nutrition among children below two years of age, the aged (65 years and above without productive capacity) and people with a severe disability;
- To increase access to health care services among children below five years of age, the aged (65 years and above without productive capacity) and people with a severe disability;
- To increase basic school enrolment, attendance and retention of recipient children between five and 15 years of age; and
- To facilitate access to complementary services among recipient households.

Figure 8 below captures the different stages of the implementation process from inputs to impact. It identifies the key stages involved for LEAP to reach its objective of reduction in extreme poverty. This results framework is the outcome of a series of discussions held with the LEAP management team and forms the starting point for the development of the M&E system.

Figure 8: Results framework for the Ghana LEAP programme



Box 4: Continued

The four key components of the theory of change are described below:

Inputs are the resources required to undertake the intervention's activities. In the context of LEAP, inputs relate to numbers and capacity of staff and flow of funds across the different tiers of implementation agencies. Programme funders and implementers have direct responsibility for acquisition, control and use of inputs; therefore, reaching the optimal benefit of resources is strictly an internal issue. These inputs are to be used for activities to achieve a range of outputs.

Outputs are the direct evidence of an intervention's activities. The output results are a function of the actions of programme implementers, and are therefore within their control. In the context of LEAP, effective service delivery relies on seven core output dimensions: 1) effective targeting; 2) payment and case management systems; 3) financial management; 4) coverage; 5) the existence of linkages with other interventions; 6) the presence of effective M&E; and 7) Management and Information (MIS) systems. As outputs are under the control of the programme, its monitoring provides a direct reflection of the programme's performance.

Outcomes respond to questions relating to the effectiveness of LEAP activities and outputs. In this context, this area helps to capture the effectiveness of LEAP. There are five expected outcomes envisaged under the LEAP programme. These are: 1) improved household consumption; 2) better health status of the recipients; 3) increased school attendance;

4) observation of the other co-responsibilities and 5) increased access to complementary services. Outcomes are differentiated from impacts in terms of timeframes. Compared to impacts, outcomes require a relatively short time to assert that the intervention has led to the desired results.

Impacts refer to long-term changes in recipients' conditions. In the context of LEAP, the expected impact of LEAP's intervention is poverty reduction.

Source: Government of Ghana (2013); UNICEF (2018)

Box 5: Engendering Ghana's 'LEAP 1000' evaluation framework

In 2015, the Government of Ghana introduced a pilot called 'LEAP 1000' to include a new category of recipients: pregnant women and children under the age of 12 months, among households who meet the poverty-related criteria. LEAP 1000 provides bimonthly cash transfers and premium waivers to enroll in the National Health Insurance Scheme (NHIS). In recognition of the fact that poverty is gendered and sustainable poverty reduction cannot be achieved without addressing gender inequalities, the impact evaluation of LEAP 1000 included an examination of **gender-sensitive outcomes. LEAP 1000 was found to have the following effects related to gender equality and more positive outcomes for women:** increased savings, greater agency and decision-making, increased happiness and life satisfaction, greater social support, and reduced intimate partner violence. LEAP 1000 was first piloted in 10 districts in northern Ghana and has since been scaled up into the LEAP programme nationwide.

Source: Authors and Ghana LEAP 1000 Evaluation Team. (2018). Ghana LEAP 1000 Programme: Endline Evaluation Report. UNICEF Office of Research 2018 Florence, Italy.

Box 6: Monitoring disability assessments in South Africa

Disability assessments are required for determining eligibility in most disability-targeted programmes. However, these assessments can be complex and assessors often have insufficient guidance and training to conduct them properly. Consequently, there are complaints that disability assessments in many settings are inconsistent and can exclude some people with disability (e.g. people with less apparent impairments).

To address these concerns, South Africa has implemented some monitoring mechanisms for their disability assessments. About 20% of disability assessments, which are required for applications for the Disability Grant, are reviewed by a panel of medical experts. These experts check for agreement between their assessments and those reached by the Medical Officers within the South Africa Social Security Agency (SASSA). The expert panel's assessments do not change the decisions of SASSA, but are used to monitor trends in decision-making amongst Medical Officers and ensure they are implementing assessments according to protocol. This form of monitoring can help to ensure the established protocol is being accurately followed so that assessment decisions are less subjective. However, it is unlikely to capture challenges within the official assessment protocols that can create barriers to access for some people with disability (e.g., assessment criteria is biased against certain types of impairments; financial, attitudinal and other barriers affecting access).

Source: Kidd et. al 2019, Kidd et al. 2018

2.2 DEFINING THE INDICATORS

Each country should develop their indicators based on country/policy/programme specific information needs:

- The **policy/programme objectives**, Theory of Change, Logical Framework (see the previous Section) and specific Service Standards (see Section 4). For example, what information do I need to assess whether Input, Activity, Output or Outcome X in the Results Framework has been achieved and to what extent? What indicators enable me to measure whether I am performing in terms of my Service Standard targets?
- The **needs of different actors and stakeholders.** For example, what does Actor X want to know about the policy/ programme and for what purpose? What information does Actor X need to adequately fulfil his/her duties concerning the policy/ programme? (See also Table 1 above)
- The **functioning of key policy/programme processes.** For example, what are the main steps involved in Process X (e.g., registration and enrolment)? Are these steps implemented as intended, and why? What could go wrong at each of those steps? What information is needed to monitor each of those steps?

While different countries develop very different systems, Grosh et al (2008) advise that a comprehensive M&E System will track indicators capturing inputs, processes, outputs, intermediate and final outputs, and the performance of its programmes. In detail:

- **Inputs:** budget, staff time and other administrative resources (though very difficult to quantify and operational costs are often not broken down by the type of activities staff engage in). Accessibility audits of programme facilities, procedures and systems are also required.
- **Output:** number of recipients, typology and profile of recipients, number of transfers and other services provided to them.
- **Outcome/Impact:** Indicators to measure improvement in recipients' consumption, incomes, wages, accessibility, empowerment, social support, etc. (depending on programme ToC) and satisfaction with the programme. Programmes may also assess the moderating factors which influence the nature, scale and magnitude of outcomes/ impacts. Note that these are difficult to collect for standard monitoring activities and pertain mostly to the realm of evaluation. Nevertheless, some form of outcome monitoring is possible (e.g., using data from national surveys).
- **Performance or efficiency** indicators, to capture the programme's cost-effectiveness. These indicators do not just focus on what was inputted or on outputs alone, but compare inputs and outputs to the goals that need to be achieved (e.g., keep costs under a certain amount, percentage of women and girls and people with disability accessing programme in comparison to the general population). Performance indicators are therefore tied to goals and objectives and serve simply as 'yardsticks' by which to measure the degree of success in goal achievement. Performance indicators are usually expressed as a rate, ratio or percentage.

- **Procurement efficiency:** whether the programme achieved value for money concerning purchases of inputs. E.g. the average cost of food procured for school feeding programmes (outputs – food procured – related to overall delivery cost)
- **The efficiency of service delivery:** how efficiently inputs were employed to produce service outputs. For example, Applications processed per staff member or per US$1,000 of administrative costs (Output – applications processed – related to staff inputs or cost). To what extent has efficient coordination and collaboration with existing partners (e.g., Organisations of People with Disabilities) been achieved? (Output – number and cost of disability-inclusive adjustments made to data collection and dissemination methodology – related to using available resources most economically to reach recipients).
- **Effectiveness:** programme's results (the change in outcomes) per unit of output. E.g., reduction in poverty gap per US$1,000 in transfers (outcome X related to the unit of output); increase in the number of people with disability accessing a service following the introduction of new adaptations for accessibility.

The selected indicators provide the **'content' of the M&E system**, (i.e., 'what' is being measured, monitored and evaluated). They should be: precise and unambiguous; appropriate to the subject at hand; available at a reasonable cost; provide a sufficient basis to assess performance; and, be amenable to independent validation (Schiavo-Campo, 1999). Indicators should provide a basis for an M&E system to **perform both functions of internal performance improvement and external accountability.** In some countries (see for example a case study on Ghana in Box 7) social protection M&E frameworks differentiate between:

- Indicators with an **operational and management focus,** that are most useful to programme managers (serving an internal performance improvement function);
- Indicators that provide an **overview of the programme's performance,** that is most useful to external stakeholders (donors, government, etc.) (serving an external accountability function).

Box 7: M&E Framework for the LEAP cash transfer project in Ghana

The Livelihood Empowerment against Poverty (LEAP) cash transfer is the flagship programme of Ghana's National Social Protection Strategy (NSPS).

The development of an M&E system is justified based on multiple objectives: a) improving LEAP operations, b) facilitating evidence-based decision-making; c) encouraging lesson learning, and d) ensuring accountability.

The availability of routine and up-to-date monitoring information should:

- allow programme managers to improve LEAP programme performance by planning the required resources to achieve the expected results of the programme;
- provide the necessary information for decision-making purposes;
- ensure that lessons learnt feed back into programme design and operations;
- apply an equity lens to ensure the programme is reaching vulnerable and marginalized groups, including people with disability, women and girls as well as the youth;
- keep stakeholders updated about the impact and effectiveness of the programme.

In this process, an implicit tension emerges between two sets of objectives. This is because programme managers have different informational needs compared to external stakeholders. The M&E framework has been designed to find a balance between these two objectives.

To respond to all these objectives at one time, the LEAP M&E framework consists of two key tools: the Core Result Framework (CRF) and the Operational Management Framework (OMF).

Box 7: Continued

Table 2 below summarises the distinction between the two tools of the M&E system.

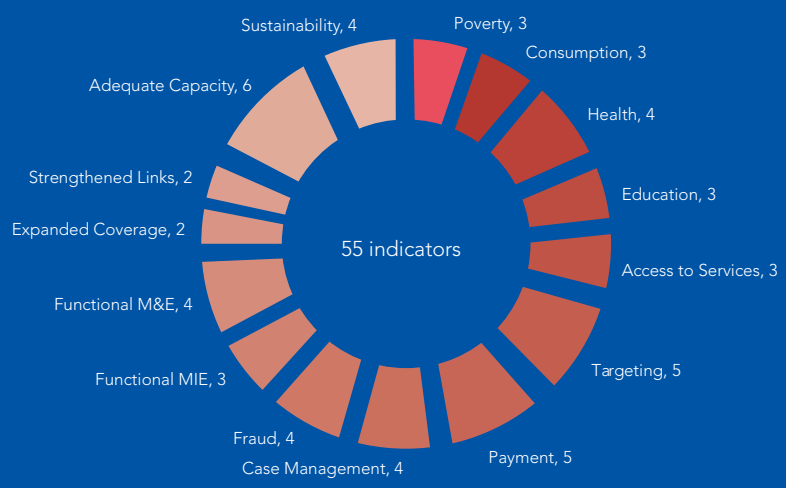
Table 2: LEAP two M&E indicators framework at a glance

DIMENSIONS	CORE RESULTSFRAMEWORK	OPERATIONAL MANAGEMENT FRAMEWORK
Purpose & Audience	Provide an overview of programme performance to external stakeholders (such as the public, government, and development partners)	Provide detailed information on programmeoperation to LEAP management
Content	Covers all areas relevant to programme performance, such as targeting, payment performance, case management and other output components, as well as impact, outcomes, and inputs (total of 15 components)	Focuses on operational procedures (Operational Manual). Provides in-depth process information for each of (eight) output components and (two) income components.
Number of Indicators	Consists of 55 indicators at the national level	200 + indicators available at the district level. Summary of 15 key management indicators
Periodically	Annual Report	Quarterly report summary of indicators at the national and regional level

The LEAP Core Result Framework (CRF) consists of 55 indicators mapped against the 15 areas of the results framework (Figure 9). These indicators measure the level of change brought about by the programme in each of the results areas. This tool provides information which enables the general public, civil society, various stakeholders, taxpayers and other sponsors to know whether or not funds have been used efficiently and effectively. In other words, the CRF is an accountability tool, and performs an outward-looking function within the M&E system. Indicators are reported every year and data is available in an annual LEAP report.

The indicators in the CRF have been ranked and prioritised jointly with the DSW for relevance to the programme and likely data availability. This list has been further revised because of the accessibility of data, internal consistency, and indicators already measured in the impact evaluation studies and by a proposed Independent Monitoring exercise. The distribution of indicators across components is shown in Figure 9 below.

Figure 9: LEAP Core results framework:



Box 7: Continued

The distribution of indicators per component

The LEAP Operational Management Framework (OMF) consists of more than 200 indicators which enable programme implementers to monitor a range of operational processes as set out in the LEAP Operational Manual. For this reason, the indicators relate to the eight output and the two input areas of the results framework. This set of 200 plus indicators will be generated periodically as a national-level aggregate. Given the wealth of information (many indicators can be considered for each district), a more manageable subset of key management indicators has been identified to provide easy access to the vast set of information. It will provide regional and district-level operational performance information. Looking at these will generate questions (why is this region doing better than that region?). These questions can then be explored by investigating indicators in the OMF that are related to the question and the region. In this way, the key management indicators help to define which information should be gathered from the OMF.

The OMF will enable programme managers to get more regular and more detailed information about LEAP's day-to-day performance to facilitate organisational learning. The OMF is therefore an operational management tool, which facilitates an inward-looking function within the M&E system.

Many of these indicators can be disaggregated to the regional and district level, and in some cases even the community level. These indicators will be run periodically at the national level, to shed light on LEAP performance across the eight output and two input components. A summary sheet with the key management indicators has been developed to enable managers to obtain the key information quickly. This should make the vast set of indicators and information contained in the OMF more accessible. It provides an entry point for managers to spot issues that need further exploration. It then allows the possibility of delving back into the larger set of indicators and information on a district level for more information about certain implementation issues if needed.

For full information on the scope and characteristics of indicators included in the LEAP strategic and management M&E framework you can consult detailed documentation available here:

Certain elements of the LEAP M&E framework are available to the general public as part of an online dashboard that can be consulted at: https://leap.mogcsp.gov.gh/

Source: Government of Ghana (2013)

2.3 PRIORITISING, REFINING AND ORGANISING

Indicators should also be prioritised, refined and organised as an iterative process (an extensive mapping of information needs can lead to some indicators that are extremely large and unmanageable). This includes:

- **prioritising indicators** based on a realistic assessment of their feasibility and usefulness (e.g., during participatory workshops with all key stakeholders, including organisations representing vulnerable groups, such as Organisations of People with Disabilities, women's groups, youth, etc.)
- **refining each indicator** to make sure it fulfils the 'CREAM' and 'SMART' criteria (see more in Box 8) and can be effectively calculated. This involves mapping each indicator back to its constituting formula (numerator and denominator) and potential data source, as well as defining how often that indicator will be collected and by whom. At a minimum, indicators should be disaggregated by sex/gender, disability status and age. Other variables relevant to the programme (e.g., ethnicity, marital status, etc.) should also be captured where possible. The number of indicators should be manageable and should not overburden the staff responsible for data collection, as this may lead to institutional resistance and poor quality of data. For an example, see Table 3 below

- **organising indicators** based on their use. For example, distinguishing between those focusing on programme operations ('management' or 'operational' indicators, that could be used by managers at all levels to assess the overall functioning of the policy/programme), and those focusing on results ('analysis' or 'results' indicators, used by high-level managers to measure progress against outcomes and for external accountability). Within each of these, indicators can then be organised by process and by Log frame level (input, output, etc.). See more inTable 2 above.

A combination of quantitative and qualitative indicators should be used to allow for a comprehensive picture of progress and changes to be documented. For example, in a gender-sensitive M&E framework, quantitative indicators will be essential to measure change numerically, such as the number of women and men benefitting from a programme or service, while qualitative indicators will be useful for assessing more complex aspects related to the 'nature' or quality of changes as a result of the intervention (e.g., women's experience of increased autonomy or mobility resulting from increased access to cash transfers). Moreover, a practical plan for guiding the analysis of indicators should also be developed. For example, it is common for M&E system to collect sex-disaggregated indicators, without investing proper time and resources to conduct a gender analysis of data, which requires specialist skills and planning.

Table 3: Indicator mapping

INFO	INDICATOR 1	INDICATOR 2	INDICATOR (etc)
Name of indicator			
Justification (why needed)			
Formula for calculation (x=y/z)			
Data sources needed to calculate (source x and y)			
Institutional responsibilities (collected by X, analysed by Y, etc.)			
Periodicity (every x month)			
Levels of disaggregation			

Source: Authors

Box 8: SMART and CREAM indicators

CREAM and SMART principles are used to select good performance indicators. SMART and CREAM describe desirable properties of M&E indicators. Defining indicators that are CREAM and SMART amounts to an insurance policy, because the more precise and coherent the indicators, the better focused the measurement strategies will be.

SMART		
S	SPECIFIC	An indicator measures only the design element (output, outcome or impact) that itis intended to measure and none of the other elements in the design
M	MEASURABLE	Has the capacity to be counted, observed, analyzed, tested, or challenged
A	ATTAINABLE/ ACHIEVABLE	The indicator is achievable if the performance target accurately specifies the amountor level of what is to be measured to meet the result/outcome.
R	RELEVANT	An indicator must be relevant. It should be a valid measure of the result/ outcomeand be linked through research and professional expertise. There is no reason tocreate an indicator which does not relate to the larger outcome
T	TIMEBOUND	The indicator is attached to a time frame. The indicator should state when it will bemeasured.
CREAM		
C	CLEAR	Precise and unambiguous
R	RELEVANT	Appropriated to the subject matter at hand
E	ECONOMIC	Available at a reasonable cost
A	ADEQUATE	Able to provide a sufficient basis to assess the performance
M	MONITORABLE	Amenable to independent validation

Source: Kusek and Rist (2004)

Box 9: Prioritizing Indicators for Ghana LEAP M&E framework

In late 2010, Ghana's flagship LEAP programme initiated a process to develop a new (M&E) system with the support of development partners. The new system is built on existing workflows and data as much as possible, seeking to improve them by providing key information and at the same time reducing the burden on front-line managers, namely the district officers in charge of implementation. Developing indicators for the LEAP programme included:

- agreeing with the government and a wide range of stakeholders on a programme Theory of Change and developing a Results Framework to prioritise indicators on that basis.
- field visits to local levels of programme administration to map existing M&E practices and assess information needs at these levels. This revealed that higher-level managers were often not aware of the challenges and the decision-making needs of lower levels.
- a set of workshops within the central agency in charge of implementation to identify any gaps, prioritise indicators and develop consensus. Consultations were also extended to other ministries that were engaged in the programme prompting discussions about the role of M&E in improving linkages with other social programmes.
- tracing core business processes (e.g., registration and enrolment, payments, etc.) for each programme being monitored to make sure that each aspect of the programme implementation could be assessed within the M&E framework.
- designing indicators and targets for each business process, linking these to Service Standards.

The outcome of the original consultation process was a set of 100+ indicators, reduced to 55 following iterative negotiations within workshops (a relevant consideration in the prioritising indicators was the accessibility of the underlying data sources). This refined set of indicators was mapped back to its constituting formula (numerator and denominator) and potential data source (including defining how often that indicator would be collected and by whom), as well as the appropriate log frame level (input, output, etc.).

The main data sources and institutional arrangements for Ghana's revised M&E system built on existing processes and structures (e.g., standard reporting was made more useful by introducing feedback and benchmarking), while also introducing a temporary Excel-based MIS while the full application software was being developed and identifying capacity gaps and training needs.

Source: Authors

2.4 DEFINING MONITORING AND EVALUATION APPROACHES

There is a wide range of monitoring and evaluation approaches that can be adapted to different circumstances, at different stages of the policy implementation cycle, depending on the information needs amongst different external and internal stakeholders and the critical issues for decision.

2.4.1. Untangle the differences between monitoring and evaluation

When discussing M&E systems for Social Protection, the standard approach is to lump the two concepts of monitoring and evaluation together, without necessarily distinguishing between the very different objectives these two activities help to achieve. Table 4 below summarizes the complementarities between monitoring and evaluation approaches.

- **Monitoring** is a regular collection and analysis of programme data to track the progress of an intervention and determine whether it is on track or not. Monitoring is necessary for efficient administration and decision-making, for improving the quality of service provision, and for the dissemination of information to bolster institutional learning and accountability. The basic principles for high-quality monitoring include the following:
  - It should be simple and useful;
  - It should be timely, relevant, dependable, credible and accurate;
  - It should be participatory and inclusive of all relevant stakeholders;

- It should disaggregate data by gender, disability status, age and other indicators of vulnerability, in anticipation of identifying potential barriers faced by different groups;
- It should be flexible, reoccurring, and routinized without being rigid;
- It should be based on pragmatism and focus on improving services for citizens;
- It should be cost-effective.

- **Evaluation** is a periodic assessment of a programme to establish whether it is achieving the intended goals and objectives and how. Evaluation is necessary to increase in-depth knowledge about one or several aspects of the intervention for learning, informing decision-making processes, and enhancing legitimacy. Sometimes the term evaluation refers to assessing changes in outcomes resulting from an intervention. This is only one type of evaluation: impact evaluation (more in Section 2.4.5).

Table 4: Monitoring and Evaluation of Social Protection compared

	MONITORING	EVALUATION
Focus	Understanding and fixing programme failures and assessing the functioning of key programme processes, for better programme management (note that this can include outcome monitoring)	Determining the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability of a social protection policyor program
Utility	Aims at continuous programme improvement andaccountability	Provides information for major decisions such as starting, ceasing, expanding, or reducing a programme
Frequency	A continuous, routine activity that should be anintegral component of any programme	Infrequent undertaking (done at certain key momentsin time), if the impact evaluation baseline is before the programme starts
Breadth	Comprehensive – aimed at all aspects of programme implementation	Less comprehensive – aimed at specific aspects of a programme theory of change or implementation
Cost	Involves low annual costs, however, set-up costs can be large	Cost varies largely depending on the evaluation methodology that is used

Source: adapted by Attah et al (2014) from Burt and Hatry (2005) and Grosh et al (2008)

2.4.2 Choosing the right evaluation approach[7]

Evaluation approaches can differ largely in terms of the evaluation questions addressed (see Box 10 below) and the methodological approaches adopted. In broad strokes there are two main families of evaluations:

- **Evaluation for formative purposes,** to inform decisions about programme or policy improvement. Formative evaluation is considered useful at two levels:
  - **Programme Level:** to improve the efficiency or effectiveness of specific programme delivery mechanisms (e.g., grievance mechanism), the appropriateness of the service delivered (e.g., size of the cash transfer, nature of farming inputs, accessibility for different populations) or integration/mainstreaming of cross-cutting issues (e.g., integration of nutrition, gender or disability in a programme);
  - **Policy Level:** to improve the functioning of the social protection system (e.g., horizontal or vertical coordination), strategy/policy or implementation plan.
- **Evaluation for summative purposes,** assessing the merit and worth of a programme or a strategy, in order to inform decisions about expanding, downscaling, merging, phasing out or redesigning it.

7 This section draws largely from Pellens (2017)

Box 10: Examples of key questions according to the OECD DAC Criteria and criteria about equity, gender-equality and HRBAP

Relevance:

- Is the policy/programme consistent with the needs of its target group(s)?
- How well does the policy/programme align with the priorities of key stakeholders, such as the government?
- How appropriate is the programme strategy to achieve the programme objective?

Effectiveness

- To what extent is the programme/policy achieving the intended outcomes?
- How can programme/policy implementation be improved to deliver its outputs more according to plans?
- To what extent is the programme/policy reaching its intended target population(s)? Amongst the target population, are vulnerable groups equitably reached (e.g., people with disability, women and girls)?

Efficiency

- To what extent does the programme acquire/use inputs (human and material resources) of appropriate quality and quantity at the lowest possible cost?
- To what extent are maximum outputs produced for any given set of inputs?
- Can the costs of the policy/programme be justified by the results?
- To what extent has the programme delivered outputs in time?

Impact

- To what extent can a specific impact be attributed to the programme/policy?
- How did the programme/policy make a difference in the lives of the target population? Were there differences in impact amongst the target population (e.g., women vs men, people with vs people without disabilities, youth)?
- What are unintended, positive or negative, outcomes of the policy/programme?

Sustainability

- To what extent will changes produced by the policy/programme be maintained?

Criteria about equity, gender equality and taking a human rights-based approach to programming (HRBAP)Equity

- To what extent did the programme reach marginalised and vulnerable groups and achieve fair and equal outcomes for them?

Gender equality

- How well did the programme promote the equal rights of women and girls and support their full participation in the political, social and economic development of their communities?
- What were the barriers and enablers that differentiated between successful and unsuccessful programme implementation and results for women and girls?

HRBAP

- To what extent did the programme achieve normativity, non-discrimination, participation, transparency, and accountability

How valuable were the results to service providers, clients, the community and/or organizations involved?

Source: Pellens (2017), Peersman (2014)

Table 5 below discusses how different types of evaluation could be relevant and appropriate in different circumstances.

Table 5: Alternative Evaluation Approaches

EVALUATION TYPES	WHEN TO USE	WHAT IT SHOWS	WHY IT IS USEFUL
Formative Evaluation	• During the development of a new programme • When modifying an existing programme	• Whether the proposed programme elements will be understood and accepted by the pop • The extent to which an evaluation is possible based on the goals and objectives	• Allows modification to the plan before implementation • Maximises the likelihood of success of the programme
Process Evaluation	• As soon as programme implementation begins • During an operation of an existing programme	• How well the programme is working • The extent to which the programme is being implemented as designed • Factors which moderate the delivery of the programme	• Provides an early warning for any problems that may occur • Provides insight on implementation processes and how they can be improved • Provides insight if programme failures are the result of design flaws or delivery gaps
Economic Evaluation: Cost Analysis Cost Benefit Analysis Cost Unit Analysis	• At the beginning of the programme (ex-ante) • During the operation of the programme	• What resources are being used and their costs (direct and indirect) compared to the outcomes	• Provides ex-ante considerations as to whether the intervention is worth undertaking • Provides managers with a way to assess the costs relative to the results and improve performances
Impact Evaluation	• During the operation of a programme at appropriate intervals • At the end of the programme	• The degree to which the programme meets its ultimate goal e.g., reduction of food insecurity	• Provides evidence on "what works" in a specific context. • Proves the goodness of a concept/design.
Systematic Reviews	• When the evidence about a given intervention • is available from impact evaluation studies across a wide range of contexts	• Gives a conclusion on a research question that was contested by summarizing evidence from all available studies	• Provides evidence on "what works" across a variety of contexts. • Allows to compare results across countries in a systematic way

Source: Adapted from Boaz et al. (2002)

While the purpose sets out the rationale for an evaluation, the objectives and scope define what the evaluation should focus on and which elements of the policy evaluation need to cover (and which not). The focus of the evaluation could be the policy as a whole, a specific programme, a pillar or strategy, or a theme (e.g., gender, disability). Objectives and scope can be discussed with stakeholders at the same time as purpose and use are reviewed during a scoping exercise. Prioritisation will probably be necessary to establish a feasible evaluation agenda. Box 11 below presents criteria that can be used to decide which programmes or elements of a social protection system to prioritise for evaluation.

Box 11: Prioritisation criteria to support decision-making on evaluation focus

- **Utility:** To what extent will the findings feed into policy decision making or programme improvement?
- **Evidence base:** Does the evaluation address an evidence gap? Is there need for addition evidence or does sufficient evidence already exist?
- **Strategic importance in terms of funding:** How important is the programme for the policy in terms of funding that it absorbs?
- **Strategic importance in terms of objectives:** How important is the programme for the policy in terms of contributing to the policy's objectives?
- **Evaluability:** Is it feasible to evaluate the programme/policy? Evaluability is looks at:
    1. adequacy of the policy/programme design,
    2. availability of information to answer evaluation questions, and
    3. conduciveness of the context to implement an evaluation.[2]
- **Timeliness:** Can the evaluation be conducted in time for the findings to be useful for decision making and action? Have programmes been implemented during a sufficient time period to be able to answer the evaluation questions?

Source: Pellens (2017)

A case study presented in Box 12 describes the approach adopted for the development of the M&E framework for the Zambia national social protection policy. The remainder of the section discusses in more detail specific monitoring and evaluation approach: participatory monitoring, spot checks, and impact evaluation.

Box 12: Case Study - Monitoring and Evaluation Framework from the Zambia National Social Protection Policy

In June 2014, the Government of the Republic of Zambia (GRZ) passed the National Social Protection Policy (NSPP). The NSPP defines SP as "policies and practices that protect and promote livelihoods and welfare of people suffering from critical levels of poverty and deprivation and/or are vulnerable to risks and shocks." The policy seeks the realization of a comprehensive and integrated SP system clustered around four pillars: social assistance; social security and social health insurance; livelihood and empowerment; and protection. In addition, the NSPP identifies a cross-cutting pillar focused on the challenges of people living with a disability (PWDs).

The policy underscores the role of monitoring and evaluation (M&E) systems in creating a social protection sector which is integrated and allows for operational synergies and complementarities between programmes. Indeed, the NSPP acknowledges that the absence of a comprehensive and robust M&E system has perpetuated the implementation of costly and ineffective programmes with limited demonstrated evidence of poverty impacts.

The NSSP M&E framework includes three components:

- **A results framework** represents the development hypothesis or theory about how an intended change will occur. It is based on a causal argument that says 'if we do a, b, and c it will lead to x, y, and z'. There are many ways to design a results framework and a variety of terminology that can be used to describe the path of change. One common way of describing this type of change in the SP sector is to define levels of results in the form of a visual representation which describes how inputs (resources) and activities of stakeholders will lead to specific short-term achievements called outputs, which in turn lead to longer-term outcomes and ultimately to the programme or policy's key goal or intended impact.
- **A monitoring framework** which expresses the results and intervention processes that will be regularly monitored and potentially evaluated. It defines indicators that will be used to measure these results and processes and includes information on how the evidence will be collected, frequency, responsibility and targets. The monitoring framework is based on the results framework.
- **An evaluation plan** which identifies 'what' will be evaluated via the formulation of evaluation questions. The evaluation plan defines the scope and focus of a more detailed evaluation and also uses the results framework as a guide.
- There are many uses for an M&E framework such as planning, communication, the inclusion of vulnerable groups (e.g., gender diverse, people with disability), consensus building, learning, management, and evaluation. In the case of the NSPP, we identified three priority uses which are to:
- **provide structure** around what the GRZ wants to achieve within the social protection sector (as defined by the NSPP), aiding in the planning and implementation of SP programmes;
- **monitor and evaluate** progress towards the implementation of the NSPP;
- **assist with decision-making** and building of consensus, coordination, and ownership around SP policies and programmes via the reporting coming out of the system.

The results framework

The NSPP states what it wants to achieve: "To contribute to the wellbeing of all Zambians by ensuring that vulnerable people have sufficient income security to meet basic needs and protection from the worst impacts of risks and shocks." The specific policy objectives and measures identified in the policy together with the policy's guiding principles and implementation plan provide insight into the intermediary steps of change (outputs and outcomes). The roadmap of change is described in the results framework as depicted in Figure 10.

Box 12: Continued

Figure 10: Zambia NSSP Results Framework

Box 12: Continued

2) The monitoring framework

For each level of the results framework, a series of possible indicators was proposed: this included the indicator description, the means of verification or how and where to acquire the evidence to measure the result, possible disaggregation, and the proposed frequency of data collection.

The structure of the monitoring framework is provided in Table 5. The full detail of the proposed monitoring framework can be found here - https://www.dropbox.com/sh/bnyyqhqkk9ph07i/AAAgbiqfH_liUXJT4TAOTm1va?dl=0

Table 6: NSPP proposed impact-level indicators

Attribute	Indicator	Indicator Name	Means of Verification	Data Collection Frequency	Disaggregation	Baseline Value	Notes

3)

The preparation of the NSSP M&E framework also comprised an evaluation scoping exercise. This is expected to result in the establishment of a rolling NSPP evaluation agenda which identifies a pipeline of evaluations for the next years and is reviewed annually after a stakeholder discussion about the purpose, use, objectives and scope of the evaluation. This resulted in the proposed evaluation agenda that is described in Table 7.

Table 7: Proposal of NSSP Evaluation Agenda

EVALUATION	RATIONALE OF THE EVALUATION	EVALUABILITY
1. Formative evaluation of SCTsdelivery mechanisms	**Purpose:** improve SCTs service delivery mechanisms, such as electronic payment mechanisms or grievance mechanisms. **Rationale:**SCTs coverage expansion makes it timely to learn how mechanism function in new districts and at an expanded scale.	SCTs have relatively good data available, objectives are clear and there seems to be demand to improve delivery mechanisms
2. Formative evaluation of SCTs linkages to othersocial services(e.g., nutrition & HIV)	**Purpose:** improve SCTs linkages with other social services, assess the assumption that through stronger service linkages the household level impact can be improved. **Rationale:** creating systemic linkages with other services may increase the impact ofthe SCTs	
3. Impact evaluation of GEWEL's Supporting Women's Livelihoods (SWL) component	**Purpose:** mix of formative, summative & knowledge focused evaluation. **Rationale:** can add to the growing internationalevidence base on SP graduation programmes	Evaluability seems adequate to the extent that primary data collection is funded
4. Evaluation of the Food Security Pack	**Purpose**: summative evaluation to decide on a redesign, continuation or phase out. **Rationale:** FSP has been in existencefor a long time without evaluative evidence about results	Evaluability would need to be assessed in terms of information availability & resources
5. Formative evaluation of the NSPP coordination(or NSPCS)	**Purpose:** improve coordination/synergy function of NSPP (or improve performance of the NSPCS) **Rationale:** Anecdotal evidence that coordination is not functioning. The new coordination strategy may require adjustment after the initial implementation.	Systematic & objective assessmentcan add value.

Source: Simon et al (2016)

2.4.3 Participatory monitoring approaches

Several approaches have been used to enhance civic engagement and incorporate the views of programme recipients into programme monitoring and design, promoting transparency and accountability. For example, in South Africa, the Department of Performance, Monitoring and Evaluation in the Presidency (DPME) has initiated a 'Framework for Strengthening Citizen- Government Partnerships for Monitoring Frontline Service Delivery', involving a Citizen-Based Monitoring (CBM) Pilot. 'This Government initiated accountability mechanism represents an effort to include citizens' experience of service delivery into their overall monitoring, evaluation and performance frameworks.' The most common tools to carry out participatory monitoring include:

- **Citizen Report Cards and Recipient Satisfaction Surveys:** participatory surveys that provide quantitative feedback on user perceptions of the quality, adequacy and efficiency of public services. They go beyond just being a data collection exercise to being an instrument to exact public accountability through the extensive media coverage and civil society advocacy that accompanies the process. To ensure a truly participatory and inclusive process, it is important that vulnerable groups, such as people with disability, women and girls, are not only included in the data collection, but also consulted in the development of the surveys. Consultations need to ensure that the questions asked on the surveys are relevant and accessible to these groups. For example, people with disability may want to include questions about accessibility, relevance and adequacy of benefits, experiences of discrimination and also provide guidance on the appropriate accessibility formats for data collections (e.g., Braille, audio, sign language, etc.).
- **Community Score Cards:** qualitative monitoring tools that are used for local-level monitoring and performance evaluation of services. These methods are particularly important for groups that face barriers to participating in research, such as literacy and physical impairments. To increase participation, relevance and empowerment, vulnerable groups such as people with disability, women and girls, need to be trained on how to use tools like Community Score Cards and also participate in data collection. These processes also need to be accessible for and inclusive of people with different types of disability. By including an interface meeting between service providers and the community that allows for immediate feedback, the process is also a strong instrument for empowerment.

Box 13: Social Protection Community Monitoring in Mozambique

The **Mozambican Civil Society Platform for Social Protection (PSCM-PS)** is a network of 35 Civil Society Organizations, whose general objective is to contribute and influence the decision-making processes that can make Social Protection services accessible to Mozambican citizens, but especially to the most vulnerable population groups.

One of the main activities developed by PSCM-PS is **Independent Community Monitoring (MCI) to the Basic Social Subsidy Programme (PSSB).** This MCI is being carried out in partnership with the Government through the National Institute of Social Action (INAS) at the national and local levels, with the aim of contributing, promoting the voice of the recipients and improving the quality and impact of the programme.

The methodology used, with the use of payroll of INAS is composed of 3 tools that help in a simple way:

Listen to the recipients through the **Community Citizen Report Card - Tool 1** (F1) with 13 questions about the functioning of the PSSB that are made to the recipient individually and anonymously;

Identification of the main problems through the **Community Scorecard - Tool 2** (F2) where the results obtained with F1 are compiled in the different communities where the survey was carried out;

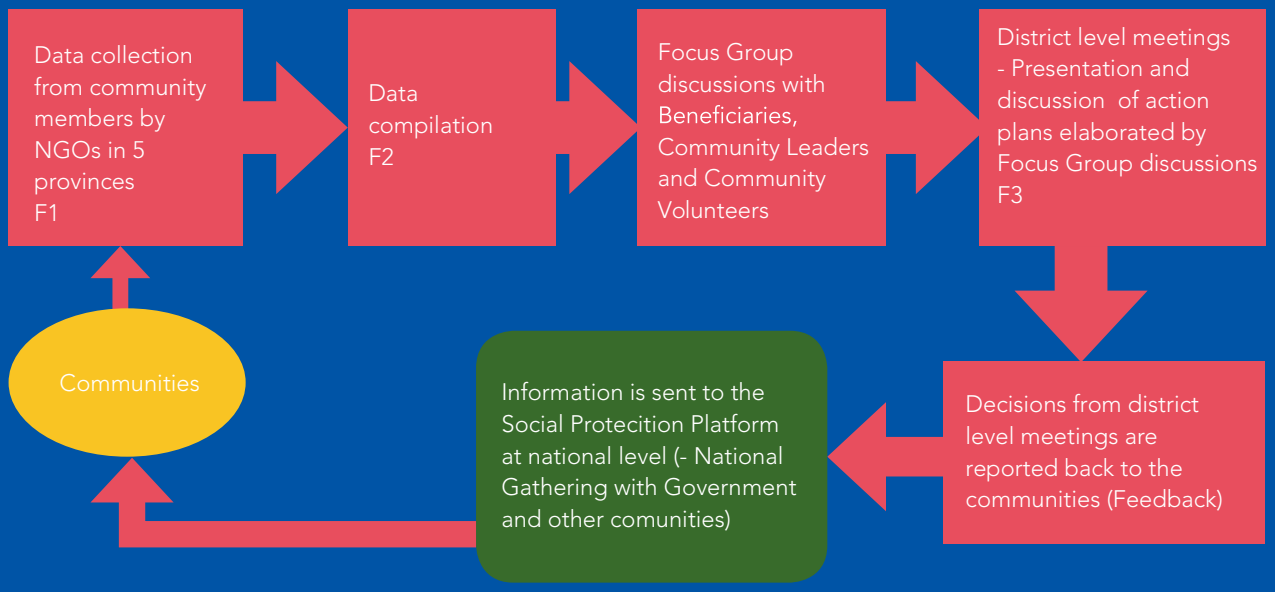
Looking for solutions at the local level with the **Focus Groups discussion** (composed of recipients of the programme and one with the community and permanent leaders) where the 3 questions with the highest negative and 3 positives of F1 are selected; and finally

Developing an **Action Plan** - the action plan is discussed with the community, leaders and community volunteers at the community level in the focus groups, and presented at a District Meeting with key stakeholders in social protection at the district level.

Box 13: Continued

This monitoring mechanism allows the incorporation of proposals for solutions to the various problems encountered in communities coming from the recipients themselves. It promotes citizens' participation and responsibility in the improvement of this public service, creates a space for dialogue among government officials and facilitates advocacy for the increase of quality and impact of service provision and is considered by the Government of Mozambique as an instrument for improving the quality of the services provided whose implementation is at the national level.

Figure 11: Community Monitoring Cycle



Source: Mozambican Civil Society Platform for Social Protection (PSCM-PS)

2.4.4 Spot Checks or independent monitoring checks[8]

Independent monitoring checks (IMCs) are a way to randomly verify the data generated at the decentralized level on social protection implementation (e.g., district payment forms, case management, and district quarterly forms). IMCs provide a structured methodology to ensure that programme operational procedures are being followed in the field. The data from the IMCs can also be used to supplement and triangulate other data. The IMC instruments are specifically designed to check field compliance and programme procedures at the institutional and household level. These instruments provide control and follow-up actions for all programme processes.

The IMC instruments are in the form of a series of yes/no questions that are easy and fast to administer (approximately 10-15 minutes at the most) and easy to enter into a database for analysis. These checks cover a range of themes, such as payment, perception of services, registration, programme understanding, enrolment, complaint procedure etc. Different IMC instruments address different programme stakeholders, for example, community members, district welfare officers, community and civil society organisations and service providers.

The data collected and calculated from the IMC process can be used to inform a range of output indicators (targeting, payment, case management, fraud) and some outcome indicators (food security, complementary services). See for example the areas covered by different proposed IMC instruments in the case of the LEAP programme in Ghana in Table 8 below. An example of the IMC process as well as copies of the forms used in the case of the LEAP programme in Ghana can be found here.

8 This section draws largely from the Government of Ghana (2013)

Table 8: Focus of proposed Spot Checks instruments in the case of the Ghana LEAP programme

THEME	ADDRESSED IN:			
	HOUSEHOLD INSTRUMENT	CLIC INSTRUMENT	DSWO INSTRUMENT	PSP INSTRUMENT
Capacity		X	X	X
Eligibility	X			
Programme Fairness/Corruption	X			
Programme Understanding of Eligibility Rules	X	X	X	
Programme Understanding of Exit & Graduation rules	X	X	X	
Enrolment	X	X	X	
Payment Process	X	X	X	X
Payment Site	X	X		X
Case Management (updates)	X	X	X	
Complaints and Grievances	X	X	X	
Programme Perceptions	X	X		

Source: Government of Ghana (2013)

Independent monitoring checks should be administered and results reported by an external organization. The argument to use an external organization is twofold. First, an outside organization provides needed objectivity and perspective as it is not submerged in the day-to-day operations of the programme. Second, one reason cash transfers programme work effectively is often the strong relationship between the national office and the regional and district offices. Random and unanticipated visits by managers can threaten and degrade these relationships to the detriment of the programme as a whole. The locations for the checks should be randomly selected to provide a sufficient sample size to provide results for annual reporting.

2.4.5 Impact Evaluation

I**mpact evaluation** aims at assessing the changes that can be attributed to a particular intervention. The objective of an impact evaluation is to ascertain the extent to which a given intervention contributes to a change in the behaviour and state of its recipients (impact and outcome level), or in other words, the impact evaluation is to estimate the **causal effect** of the intervention on a given outcome variable.

The term causal effect is key to understanding the key focus of an impact evaluation. Causal effect means that some change has happened, and s**uch change can be attributed to a specific intervention**, and not to any other factor, initiative or characteristic of the context. The whole purpose of an impact evaluation is to isolate change that is due to the intervention from change that may be due to other "confounders".

Key to the analysis of a causal effect is the concept of **counterfactual.** The counterfactual represents how participants would have performed in the absence of the intervention. In a situation where recipients of a universal pension are having three meals per day, the question one would ask is how many meals would the recipients have been having, had they not been receiving the pension?

- Assessing the impact of an intervention by comparing what happened to the recipients (the traditional before-after analysis) can be very misleading as the change observed may not be due to the intervention but to other factors.
- The question an impact evaluation aims to answer is not "what happened with the recipients" but rather "what would have happened if the recipients had not participated in the intervention". This is the so called "counterfactual".
- The problem is that it is not possible to directly observe the "counterfactual", as it is not possible to observe simultaneously a recipient receiving and not receiving the intervention.

A counterfactual can be found among the non-recipients in the form of a control group. It is not difficult to find a control group, what is difficult is to find a good **control group.** The goodness of a control group depends on its comparability with the recipients (treatment group). If the two groups are not comparable the evaluation can mistake differences in characteristics between the groups for the real (causal) impact of the programme. For instance, in a district with cash transfers, one could have the control group drawn from households with higher literacy levels because they live in an area serviced by a school while the recipient group is from an area with no school facilities. The evaluation would find that the control group had higher literacy levels, suggesting the cash transfer had negative impacts on schooling, when in fact the difference has nothing to do with the programme.

Much of the challenges with impact evaluation design have to do with ensuring that control and treatment groups are comparable. To eliminate differences due to the selection process, the ideal approach is to allocate units to the treatment or the control group based on a random allocation process (lottery). By design (due to the law of large numbers) the treatment and control groups will be fully comparable (in both observable and unobservable characteristics). This is the approach used for so called **Randomized Controlled Trials (RCT)** which are commonly used in the medical field (i.e., to test the effectiveness of new drugs) and have increasingly been used to test also the impact of public policies, including in the field of social protection.

RCTs are not always feasible and there are alternatives, with varying degrees of reliability in providing credible estimates of the real "causal effect", as discussed in Table 9. For more reference on impact evaluation approaches and methodologies, you can consult the 3ie website: https://www.3ieimpact.org/en/evaluation/resources/impact-evaluation-resources

Table 9: Impact Evaluation Designs

DESIGN	KEY FEATURE	TECHNIQUES	WHEN TO USE
Experimental Design	Randomly assigned intervention and control groups	Randomised Control Trial	When you want comparability of treatment and control groups by design
Quasi-Experimental	Quasi-experimental designs identify a comparison group that is as similar as possible to the treatment group in terms of baseline (pre-intervention) characteristics.	There are different techniques for creating a valid comparison group, for example, regression discontinuity design (RDD)and propensity score matching (PSM),	Where ethical, political or logistical constraints, likethe need for a phased geographical roll-out, rule out randomisation.
Non-Experimental	Not statistically matched groups or groups compared to itself	Often takes the form of ex-ante ex-post comparison for a group of recipients only, with no consideration of the "counterfactual"	When there was no control group. Causal attribution is in this case very difficult to determine through quantitative methods.

Source: adapted from H. White and S. Sabarwal (2014)

Box 14: Are Impact Evaluations the Silver Bullet?

Impact evaluations have become extremely popular in the development field (see figure below), and particularly in the social protection field. Practically all major social cash transfer programmes in southern and eastern Africa have been subject to an impact evaluation during the last 10 years and this has contributed in many cases to their coverage expansion and establishment as national programmes (see for example Box 1 for Zambia).

Yet impact evaluation are only one of the various M&E tools and approaches that are available (see table 4 above). Impact evaluations perform extremely well in regards to the specific question they aim to answer: the causal effects of an interventions. Depending on circumstances such a question may be appropriate, or too narrow.

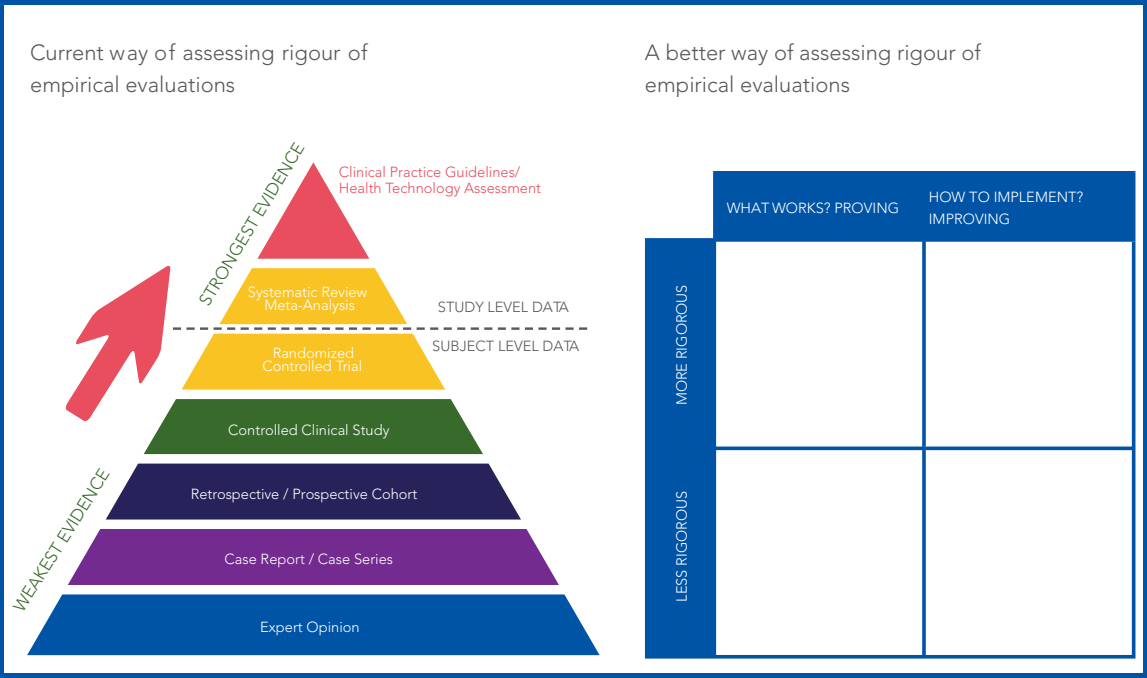
In considering whether an impact evaluation is the appropriate instrument in a given circumstance, the following limitations of the impact evaluation approach have to be considered:

- Impact evaluations are generally not equipped to determine circumstances under which an observed impact could be replicated, other than the specific ones under which the study took place (often called problem of external validity). They are not based on any theory that can be used to model and anticipate impacts ex-ante, but can effectively determine impact only ex-post;
- Impact evaluations can be expensive as they require most of the time collection of dedicated information at repeated points in time for both treatment and control groups;
    - Impact Evaluation may be considered unethical or unfeasible under certain circumstances. They may not be accepted by policy makers and the target population who may resist the idea (or it may not be technically possible) that a programme withholds benefits from a certain group (the control group) in spite of the deep pressing needs to receive support. Conversely gradual expansion of social protection programmes is often a necessity as governments don't have capacity to expand country wide at the same time: control groups can be effectively identified as part of a staggered programme roll-out plan (those areas or individuals who will enter the project at a later stage);
    - Impact Evaluations are often deemed as being "black box" evaluations, producing very rigorous evidence on what was the change under given circumstances, but not sufficient understanding of the process leading to that change and the critical implementation factors that led to that change. Impact evaluation does not necessarily explain the why and how the result has (or has not) been achieved. Impact evaluations are more versed to "proving" the goodness of a given intervention, rather than providing learning so as to "improving" the design or implementation of such intervention. Therefore, impact evaluations are increasingly combined with process evaluations to understand the role of programme operations and how these contribute (or undermine) programme results and impacts.
    - The collection of qualitative data as part of process evaluations may also offer deeper insights into the situation of vulnerable groups, such as people with disability, whose experiences may not be sufficiently captured in impact data as sample sizes for this population tend to be too small for this type of analysis. This is more the case when needing to know the experiences of persons with different types of disability – it is rare to have a large enough sample size to be able to disaggregate impacts on different disability groups (see Box 14 for further information on common pitfalls in measuring disability in evaluations). Furthermore, outcome and impact indicators often do not capture the nuances in experiences of vulnerable groups and qualitative data can be useful in helping to design context-appropriate responses.

Rigour stands as the principal standard empirical evaluations have to fulfill. However, rigour is sometime only defined in relation to proving what works. The left panel of Figure 12 below describes empirical evaluation according to the rigour of the evidence they provide on whether a given intervention will achieve a particular impact. RCTs clearly stand as the most rigorous approach to evaluation.

Box 14: Continued

Yet assessing the impact (doing the right things) is not always what policy makers need. Questions of rigour similarly apply to other kinds of empirical evaluation that focus at generating evidence to improve implementation (doing things right).



Source: Kleinfeld (2015), Deaton and Cartwright (2016)

While OECD good practice principles in impact evaluations contain inclusiveness (e.g., gender equity and disability in particular) as cross-cutting criteria, some impact evaluations are designed to focus specifically on measuring gender or disability-related changes (either as stand-alone studies or including these topics within overall impact evaluations). Box 15 describes how disability was incorporated in the impact evaluation of the Child Grant Programme in Lesotho and Box 16 provides examples of the evaluation of gender-specific outcomes in Zambia's Child Grant Programme.

Box 15: Impact evaluation of the Child Grant Programme in Lesotho and lessons learnt for disaggregation by disability

The Child Grant Programme in Lesotho is an unconditional cash transfer that is targeted to poor households with at least one child. The programme originally began as a pilot programe in 2009, which was then expanded nationally by 2015. During the expansion, a community Randomised Control Trial was conducted to establish the programme's impact. The research team worked with the Lesotho government to randomly decide which areas the cash transfer programme would be rolled out to through a public lottery. Then, households receiving the Child Grant Programme in the roll-out areas were compared to similar households in the non-roll out areas, who then did not receive the Child Grant Programme (control group). The impact evaluation found several positive impacts of the Child Grant Programme for recipient children and their households. For example, there were improvements in schooling outcomes for children and a reduction in household food insecurity.

A recent study disaggregated the impact of the Child Grant Programme by disability status. It found some modest improvements in recipient households with members with disability in areas such as health status and access to care, decreased food insecurity and increased participation in waged employment. The impact of the programme was in several cases greater for recipients with disability compared to recipients without disabilities (e.g., greater increase in engagement in paid work amongst adult recipients with disability compared to adult recipients without disability). However, recipients with disability were still worse off compared to people without disabilities even after the receipt of the cash transfer on many indicators (e.g. adult recipients with disabilities still much less likely to be engaged in paid work compared to adult recipients without disabilities).

The disaggregation of this impact evaluation by disability also revealed several challenges, which need to be considered and addressed during the study design of future impact evaluations. Importantly, even though the overall sample size was large, the sample of people with disability was small. Consequently, some analyses – particularly amongst children – were either not possible or may have underestimated the impact of the programme. It was also not possible to disaggregate findings further amongst people with disability (e.g., by gender, disability type), which can be important for tailoring policy responses. Second, the way in which disability was measured in the study is not in line with international recommendations, which likely lead to the underestimation of disability prevalence. This underestimation both affected sample size and may have skewed findings to representing the situation of people with more severe and only certain types of disabilities. Finally, some of the indicators used may have not fully captured areas of concern for people with disabilities. For example, the study included indicators on school enrolment, but not on if the school offered inclusive education.

When designing an impact evaluation, researchers and other implementers should consider the following to ensure that results can be disaggregated by and relevant to people with disability:

Disability is measured using validated, recognised tools such as the Washington Group Questions

Sample sizes are large enough to disaggregate by disability status and other characteristics of interest. Sample sizes may need to be increased overall or through purposive selection of people with disability (e.g.,through their representative organisations and local leaders).

Design indicators that are relevant to people with disability. Consultations with persons with disability and their representative organisations in the study area can help in identifying appropriate indicators.

Sources: de Groot et al 2021, Tiwari et al 2016

Box 16: Mixed-method evaluation of gender impacts of Zambia's CGP

UNICEF commissioned a mixed-methods evaluation of the Child Grant Programme (CGP) in Zambia to assess the impacts of its cash grant on a wide range of well-being indicators, including women's empowerment. The quantitative research component included a four-year longitudinal study among 2,519 households in three rural districts. Four follow-up surveys on the households were collected after 24, 30, 36, and 48 months. The survey instrument contained questions addressed to the principal female recipient in each household on women's economic empowerment and economic security, among other topics. Survey items included the amount of women's cash savings, the operation of small businesses, intrahousehold decision- making and future aspirations.

The quantitative surveys were followed by qualitative studies to better understand the findings. A one-time series of in-depth interviews was conducted in Kaputa District after the conclusion of the survey. Thirty in-depth interviews with women and ten in-depth interviews with male partners were conducted to find out why and how the changes observed in the survey took place, and determine in greater detail how the programme affected women's decision-making and overall intrahousehold dynamics. Key informant interviews with non-recipients were also conducted to examine any underlying differences in their responses compared with those of recipients. The qualitative research helped gain deeper insights into how women and men viewed and conceptualized 'empowerment' in their local settings and the impacts of the CGP on these outcomes.

Source: Bonilla et al, 2017 and FAO, 2018.

2.5 DEFINING THE DATA SOURCES

Primary data collection for M&E can be an extremely time-consuming and costly activity. For these reasons, identifying existing data sources (i.e., 'where' data comes from), establishing their usefulness for M&E purposes, and planning them carefully to deliver exactly the indicators needed is an important task.

So what are the most useful data sources for a social protection M&E system? Table 10 below outlines the potential strengths and weaknesses of the main data sources available for this sector. These need to be assessed against the country context (e.g., institutional set-up, existing databases, etc.).

Table 10: Strengths and weaknesses of selected data sources for a Social protection M&E system

	STRENGTHS	WEAKNESSES
Programme Management Information Systems (MIS) (see also ModuleMIS)	• Routinely and readily available • Large sample size (e.g., all households registered) • Low cost • Easy to add additional reports and potentially import da ta from other sources • Can generate useful performance indicators • Allows longitudinal tracking	• Cost/time of designing high-quality MISand reporting • Administrative data cannot measure all outcomes and cannot be used for inferences • Data is only available when the client is 'in the programme' • Indicators may not be specific enough to allow monitoring of vulnerable groups, such as women and girls, people with disability
Standard periodic administrative reporting from visits, spot-checks, audits, etc.	• Simple and often already in place • Can be useful to sort out ongoing implementation issues • Generates information to recipient level	• Not always filled, used and analysed (often paper-based) • Inconsistently applied and focused on procedures (reporting to the next level of hierarchy) • Not all forms and info reaches the central level in a useful way (e.g., missed payments) • Risks being 'anecdotal'
Qualitative ad-hoc studies	• Low cost due to smaller required sample size • Essential to follow up on issues raised by analysing numbers and to understand why and how things are going wrong • Can provide insights on how problems could be addressed and solved • Can examine issues in-depth and understand how the programme works for vulnerable groups	• Require capacity for design, implementation and analysis • Low sample • Risks being 'anecdotal' if low quality
Other administrative databases	• Integrating data with HR and Accounting information for performance indicators • Link with Civil Registry and other sectoral databases for integrated M&E • Bank/payment provider database for monitoring of payments	• Need to set up institutional arrangements with different stakeholders • Cost of setting up data linkages and coordination • Ideally requires Unique ID (e.g., National ID number) for linking • Ensuring data comparability • May not allow for disaggregation by disability and other indicators of vulnerability due to insufficient sample sizes

Table 10: Continued

	STRENGTHS	WEAKNESSES
Official Statistics data (Census, Household Budget Surveys,Living Standard Measurement Surveys, Labour Force Surveys)	• Can guarantee a wealth of information (household income, consumption, education, health status, etc.) at a very low cost • Comprehensive national data from Census or representative surveys could be used to calculate caseload, poverty levels, coverage, etc. • Could be used to assess the impact and targeting effectiveness if extra question on programme receipt added	• Requires building institutional relationship with national statistics office • Requires high capacity for analysis • Calculating estimates for lower administrative levels (e.g., district) may not be representative (e.g., especially for under 5) • May not allow comparison of recipients and non-recipients • Ensuring data comparability • May not allow for disaggregation by disability and other indicators of vulnerability due to insufficient sample sizes
Externally contracted impact evaluations	• Important for understanding causalrelationships and generating counterfactuals • Needed for indicators analysing targeting effectiveness, impact onconsumption, poverty, etc. • Impartiality of evaluators decreases the risk of bias	• Very high costs (require independentexternal evaluators) • Results available late in the policy process • Seldom feed into programme improvement and planning
Community Monitoring	• Very important to gain bottom-up monitoring and assess satisfaction withservice delivery (e.g., Citizen Report Cards, Perception Surveys, etc.) • Allows for the participation of community groups and civil society organisations to represent the voices of vulnerable groups	• Setting up incentives to make this happen; institutional arrangements, etc. • Costly requires capacity and could beunsustainable

Source: Authors

It is important to evaluate and select these potential data sources based on four main criteria:

- **mixing monitoring and evaluation components:** ensuring that monitoring functions do not get over-shadowed by evaluation objectives;
- **building on existing data sources:** helps to reduce cost, makes a system more sustainable, requires less active management and avoids duplication of work. Such sources can be classified as internal (generated and managed by the programme) and external (managed by external actors – require coordination). For example:
    - Internal: Programme MISs and integrated systems for information management in the social protection sector can be programmed to offer a vast array of standard M&E reports (see MODULE MIS);
    - External: Official Statistics data could offer great insights by simply adding one question on receipt of benefits to existing surveys (e.g., household budget survey);

- **ensuring triangulation** of several types of data sources, both internal and external to the programme. This can enhance analysis potential (see boxes below in this section for an example);
    - While data from programme MISs provide substantial information, they do not necessarily help to understand how and why a programme is delivering and meeting recipients' needs (or isn't). Qualitative research and approaches to participatory monitoring, triangulated with other sources, can help to address these fundamental issues;
- **minimising the burden of data collection and analysis:** ensuring the data for the system is primarily designed to be generated as an integral part of normal administration rather than as an additional task;
    - Ideally, an M&E system will feed to the largest possible extent from data that is generated in any event as part of normal operational procedures, so to minimize data collection efforts that are specific to M&E processes. For example, data are entered from the PMT questionnaire into the MIS for recipient selection and can be used for M&E without any additional effort;
    - This also implies automating reporting functions where possible (for example within a programme MIS).

    Other than the strengths and weaknesses discussed above, a country wishing to set up an M&E framework should also consider each data-sources: main uses and focus (which areas and indicators within the framework it would address); accessibility (how easy is it to use in practice, especially in the short and medium term); recommended frequency (frequency with which data from that source should be collected and analysed); sample size and; potential cost.

Capturing the experiences of vulnerable groups

Data for M&E activities must be sensitive and inclusive of the experiences of vulnerable groups, such as women and children and people with disability. Approaches to collecting more gender-sensitive and disability-inclusive data include:

- **Awareness training:** Training M&E staff to have gender and disability awareness and a general understanding of vulnerability and equity in the design of M&E frameworks and systems to design and deliver more inclusive and comprehensive evidence, and ethically collect data, protecting the identity and privacy of people participating in interviews or other data collection activities.
- **Minimum standards and accountability:** Developing minimum standards in collaboration with organisations representing vulnerable groups (e.g., women's groups, Organisations of People with Disabilities, etc.) to advise on the M&E process to ensure relevant data on the experience of vulnerable groups are captured; Publish the standards and other relevant communications in formats accessible to vulnerable groups (including Braille, audio, large print, digital text conforming with disability accessibility standards, etc.); Co-designing evaluation questions with representative groups to ensure they capture the experience of vulnerable groups (e.g., access to services that are inclusive and appropriate, addressing discrimination, unpaid care work, etc.).
- **Design and data disaggregation:** Capturing details such as enrolment, drop-outs, access to linked programmes, intra-household issues, and barriers faced by vulnerable groups in monitoring systems; Ensuring that data used for design, monitoring and evaluation allows for disaggregation by disability, gender, and other vulnerability indicators. This may include the use of specific tools to identify and capture the experiences of vulnerable groups. For example, the Washington Group question sets are recommended by the United Nations and other groups for measuring disability in surveys; Ensuring the sample size in evaluations are large enough to conduct meaningful analysis about the experiences of vulnerable groups, including people with disability and other minority groups; Ensuring data collection methods are accessible to vulnerable groups to all self-report (e.g., using Braille, audio, large print, simple language, etc. to meet the needs of participants); Ensuring that relevant household members are engaged to answer questions about their circumstances rather than relying on one household "head" to speak on their behalf.
- **Regular reports:** Produce regular reports on gender, disability, and accessibility issues (and any other issues relevant to vulnerable groups) based on information integrated into the MIS programme.

Source: Information adapted from UNPRPD, UNICEF & ILO (2021)

The proposed M&E system for the social protection programme of the National Institue for Social Accion (INAS) in Mozambique is based on three main data sources:

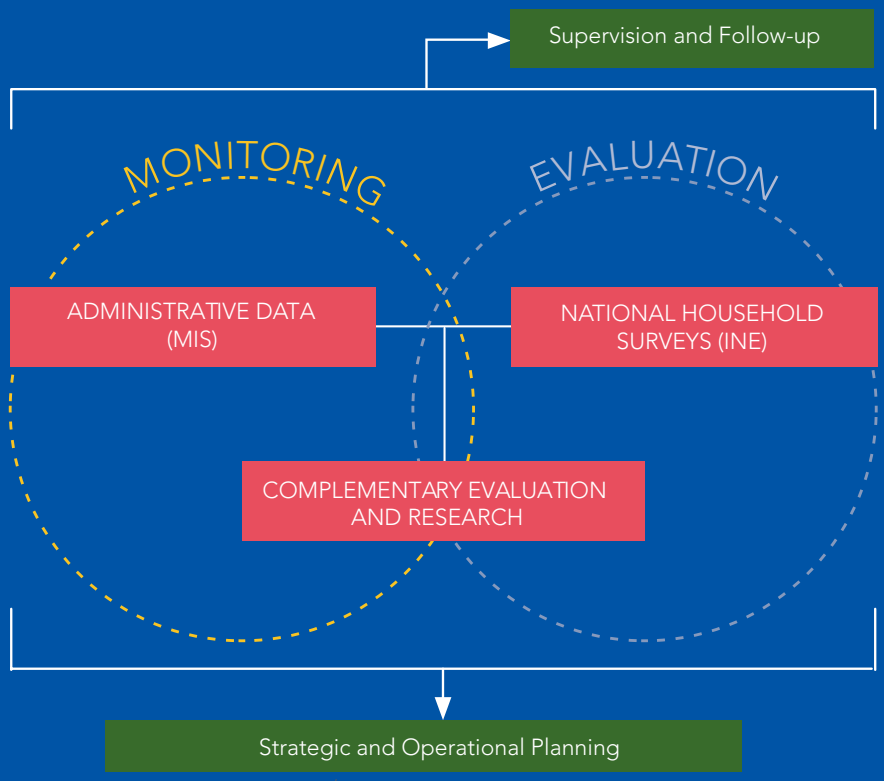
The Management and Information System (e-INAS) allows to collect, collate, and manage all information relevant at different stages of programme implementation (selection, payments, case management, etc.) in this way allowing effective and permanent monitoring of programme implementation and outputs.

The National Statistics Institute (INE) will collect and analyze information through national surveys that can be critical to perform monitoring and evaluation of INAS programmes

Complementary evaluations and surveys will be conducted by external stakeholders and INAS's staff, to provide additional insights on specific topics of interest, including the perceptions of recipients and the impact of INAS programmes. This will also include a community-based M&E system (see more on this in Box 11 above).

As reflected in Figure 13 below, the three M&E instruments will feed into two critical processes for INAS continuous improvement: a) supervision and follow-up to decentralized implementation levels, and b) strategic and operational planning.

Figure 13: The basic components of INAS M&E system



Source: Government of Mozambique (2014)

Box 18: M&E tools in a low-capacity setting

Ethiopia exemplifies the difficulties of running an M&E system in a low-capacity, low-income country and the need for continuous adaptation and simplification. Its Productive Safety Nets Programme (PSNP) has a good M&E system, designed to track progress on a range of inputs, activities, and outputs for accountability purposes as well as to allow prompt corrective action as bottlenecks are identified. The system aimed for simplicity to account for the low capacity of the programme's frontline units.

However, implementation of the monitoring plan encountered numerous logistical obstacles, with only 40 of 232 districts reporting (with delays) during the first year of programme operation—and the remainder not reporting at all. The major stumbling blocks included the lack of local staff, the poor qualifications and high turnover of existing staff, and the poor infrastructure in some districts (for example, about 20 per cent lacked electricity). To generate a minimum amount of monitoring data, a number of additional systems were put in place.

First, to assess the programme with respect to the number of recipients and actual disbursements, the programme instituted a sample-based emergency response system, where information was collected via telephone from around 80 districts on a twice-weekly basis.

Second, four- to six-person rapid response teams were formed to perform spot checks (four times a year at the federal level and eight times a year at the regional level).

Third, the Food Security Coordination Bureau instituted a system of roving audits to investigate compliance with financial rules, disbursements and payments, and appeals and complaints to provide more timely information on compliance than the normal annual auditing system.

Finally, some 80 PSNP public works projects were reviewed twice a year through spot checks to investigate the quality of planning and implementation.

In the meantime, the programme further simplified its monitoring system through such steps as halving the length of the M&E manual and invested more resources in training staff involved in M&E activities. The simplification of the M & E system and the development of a less ambitious emergency response system were appropriate responses to low capacity. Even though the formal monitoring system is now starting to show some improvement and provide more reliable data on basic programme operations, additional monitoring instruments have been kept in place, as they provide more in-depth and often more qualitative information on overall programme performance.

Source: World Bank (2010)

Box 19: Triangulating administrative and census data: South Africa's research on take-up of social assistance benefits

A good social protection system should reach the eligible, make data on take-up rates available, and undertake research to identify the factors that influence take-up and prevent accessibility. In South Africa, for example, the Taylor report investigated the effectiveness of the social grants programme. The 'Taylor Committee Report13 into the Social Security System', published in 2002, reported estimated take-up rates of 90 per cent for the Disability Grant, 85 per cent for the State Old-Age Pension and only 20 per cent for the Child Support Grant.[9]

Through effective targeting, the Department of Social Development wanted to maximize take-up among those eligible for social protection assistance. Using a combination of geo-coded administrative data on grant recipients (for the numerator) and census data (for the denominator), eligibility and take-up rates of the Old- Age Grant and Child Support Grant were calculated and mapped at municipality level to identify areas with low take-up rates.[10] [11] The findings were used by the Department of Social Development to focus its efforts on promoting take-up. Despite a significant increase of 3.79 per cent in the take-up of the Child Support Grant in 2012, SASSA partnered with UNICEF to locate the children still excluded from the programme and identify reasons for some age groups. It was found that urban and peri-urban exclusion was higher than rural exclusion[12] and resulted in a door-to-door campaign to raise awareness for the Child Grant among the relevant population before the Integrated Community Registration Outreach Programme (ICROP) visited the area.[13]

9 Taylor, 2002
10 Noble, M., Barnes, H. Wright, G and S. Noble, (2006), p. 2
11 Noble et. Al (2005)
12 Noble, M., Barnes, H. Wright, G and S. Noble, (2006), p. 20
13 SASSA (2014), p. 9

2.6 DEFINING INSTITUTIONAL ARRANGEMENTS

The institutional arrangements for Social Protection M&E should describe the roles and responsibilities of different organisations and actors within the system, while also outlining 'how' the information will be collected, compiled, analysed, reported, and used. Guiding principles include the following:

- Duplication of reporting lines, parallel or dual reporting within government structures, in some cases including parallel reporting to external funders/implementers (e.g., NGOs, Donors)
- Information extraction, limited feedback provided to decentralized levels and recipients
- Weak link to decentralized decision makers (district, province administration)
- Inward-looking M&E system
    - Limited/unstructured role for external independent stakeholders (civil society and media)
    - Weak link with external partners that undertake relevant M&E work (Statistical Office; Research institutions)
- Ensuring the institutional arrangements of the M&E system **reflect the overall institutional structure of the implementing organisations** (at all levels of decentralisation) while filling any gaps.
- Aligning all activities with the **National Planning Framework**, with guiding legislation based on the government-wide monitoring and evaluation framework.
- Working as much as possible with **existing systems, staff and processes** and helping improve them (build capacity, etc.), based on an initial Capacity Assessment (see MODULE ADM)
- Building institutional arrangements with new actors for M&E purposes: requires time, **dedication and in somecases legal frameworks or memoranda of understanding** (e.g., adding an extra question on benefit receipt to a national survey by the Statistics Bureau)
- **Explicitly budgeting for M&E activities from the outset** and thinking through related staffing needs

- Tasking one person at all levels of management (central, province, etc.) specifically with M&E duties as a priority and possibly not as an add-on to other responsibilities (creating the role of **M&E officer**);
- The main actors in a typified social protection M&E process, and their key roles, include:
- **National level** (e.g., 'Social Development Ministry'/'social welfare Ministry'): Within the relevant Ministry, a dedicated M&E Unit should have oversight of each Social Protection programme and provide strategic advice on the policy and system's overall direction. It should also evaluate the work of the institution and the quality-of-service delivery. Its overarching purpose would be to ensure efficient and effective social protection administration while fostering transparency and accountability and providing oversight and support to the cabinet and parliament

Its responsibilities include:

  - *M&E plan* development and management (formulating plan, overseeing its implementation, ensuring sufficient resources, mobilising relevant technical assistance where needed, etc.)
  - *Data collection* (coordinating supervision activities, overseeing data collection by lower levels and providing feedback, ensuring adequate data sources e.g., MIS, liaising with partner data-providing institutions, etc.)
  - *Analysis* and reporting (regularly analysing and using information collected, ensuring timely decisions are taken based on findings, signing off periodic reports by lower levels and compiling national reports, coordinating responses to queries from researchers, parliament, civil society organisations, media, etc.)
  - *Dissemination* (facilitating the wide dissemination of M&E reports and other outputs to all relevant stakeholders) in accessible formats (including Braille, audio, large print, videos with subtitles and sign language interpretation and other formats appropriate for people with disability)

- **Decentralised level** (e.g., District) At the provincial/district/decentralised level dedicated M&E officers should liaise regularly with the Ministry. Their role should be focused on overseeing the implementation of M&E activities in the field as specified in the M&E plan, and making suggestions for improvements. This includes checking the consistency of reports from lower levels and following up where discrepancies exist; ensuring timely submission of forms and standard reports to the Central level; ensuring the correct use of MIS; undertaking supervisions; managing relationships with other stakeholders at the decentralised level in relation to M&E; participating in training and ensuring training of lower levels, and; responding to any ad-hoc requests from the central level.
- **External actors** (e.g., NGOs, evaluation council/agency, academia, external evaluators, community and civil society organisations, etc.) M&E of social protection programme operations and performance is not necessarily only conducted by programme implementers. In some cases, it can be important to formally or informally involve external actors. For example:

  - NGOs may be contracted to perform community monitoring functions, ensuring external accountability.
  - A **national 'evaluation** council' may be created to ensure performance is independently evaluated by adopting a state-of-the-art methodology (e.g., in Mexico).
  - External evaluations can similarly be carried out by externally contracted consultants/researchers or academia. National academics can further be involved to support the analysis of existing data.
  - The national **Statistics Bureau** could play an important role in providing and analysing relevant data.
  - Organisations for People with Disabilities, older persons' organisations, women's groups, etc. can help reach recipients, representing their experiences and need to be consulted to ensure M&E activities correctly capture data about the experiences of vulnerable groups.

It is not common practice to perform summative evaluations in-house because of a) lack of capacity, and b) the need for an independent evaluator. Externally contracted evaluators can perform this function at a cost which increases with the level of 'rigour' of the evaluation methodology. Some countries have tackled the lack of in-house capacity and the issue of high costs of contracting-evaluations by establishing national councils mandated to conduct a sectorial evaluation for social protection.

While they are public bodies, these councils have autonomy from the rest of government and often bring together specialists with high capacity in evaluation (e.g., Mexico, see Box 18). Conversely, the participation of in-house staff in formative evaluations is critical to support the learning process (see more in Section 3.2.2 below).

Box 20: Case Study: M&E processes and institutional arrangements in Ghana

When an assessment of the overall institutional context and M&E situation for the LEAP programme was conducted in 2013, it revealed the strengths and weaknesses of the system that are common to many similar projects in the region:

- The MIS system was able to support basic operational processes. However, overall, the current nature of M&E remained highly unsystematic.
- There were no clear lines of responsibilities established for data collection, analysis, reporting, and dissemination.
- There was no consistency in reporting tools:
    - Reporting by the LEAP programme both at the district and national level was found to be too inconsistent and unstructured, with limited use of a standardised reporting format.
    - The reporting process was often triggered by payments and targeting, meaning it is not necessarily done regularly.
    - While the reports were activity-based and descriptive, they lacked some of the contents needed for analytical purposes.
- Limited feedback was provided to implementers at the decentralized level, who saw little value in reporting information.
- The capacity for undertaking M&E was generally very weak.
- There was minimal use of M&E information for decision-making.

A new M&E system was designed to help systematise M&E activities by establishing clear reporting lines, improving tools for data collection and reporting and promoting the use of M&E information for planning and decision-making. The DSW articulated clearly the ideal M&E system for the LEAP programme to be one that: reflects current reality in terms of capacity and staffing levels; is responsive and reflects the evolving nature of LEAP programme; is designed for the short term and with a plan for the longer term, where different components are progressively incorporated; is simple and easy to use yet innovative; is integrated and feeds into and used by other stakeholders and generates a sense of demand for data.

Principles of the proposed M&E system

The LEAP M&E system is based on two key principles:

- To build on existing systems: Building on existing systems helps to improve these systems and make them more sustainable, and avoids duplicating work. For example, District Social Welfare Officers were already sending quarterly report forms to the LEAP Management Unit, but not all did this and relevant information was not even asked for. Hence, a starting point for the work was to understand why not all District Social Welfare Offices, what could be done to improve this, and which information should be collected.
- To develop tools that minimise the burden of data collection: M&E requires the collection of data, which can be burdensome for those who are supposed to provide it. Hence, a successful M&E system works best if the data-gathering effort is minimised and those who gather data can see the benefits of doing so. Ideally, an M&E system feeds to the largest possible extent from data that is generated in any event as part of normal operational procedures, so to minimize data collection efforts that are specific to M&E processes. For example, data are entered from the PMT questionnaire into the MIS for the purposes of recipient selection and can be used for M&E without any additional effort.

Box 20: Continued

The three dimensions of the proposed M&E system

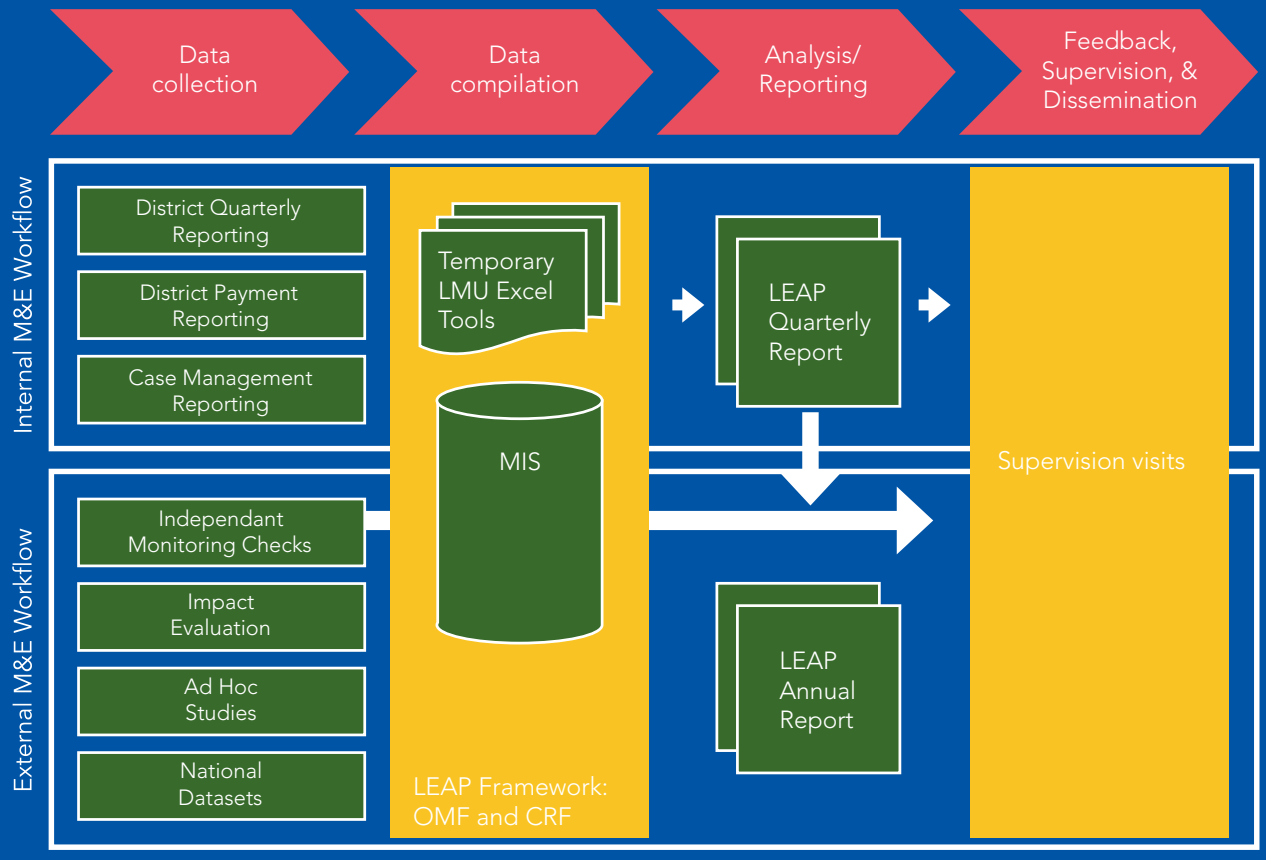
Any system that provides information on a programme's progress and success will need to state which information (indicators) need to be tracked, their data sources, and the process by which the information is used. The proposed M&E system has thus developed along three dimensions:

- The M&E framework provides the 'content' of the M&E system, namely the indicators (i.e., the 'what'). This looks at what the M&E system will measure and monitor;
- The M&E platform describes the data sources, how they relate to each other and the instruments used for data collection, compilation, analysis and reporting (i.e., the 'where'); and
- The M&E process describes how information is used as well as how the entire M&E system operates.

The latter refers to the roles and responsibilities around the different components of the M&E system, with respect to the question of how information is collected, compiled, consolidated, analysed and used (i.e., the 'how').

In summary, the M&E system of LEAP is built in Figure 14 below.

Figure 14: Overview of the LEAP M&E system



The M&E process – Reporting

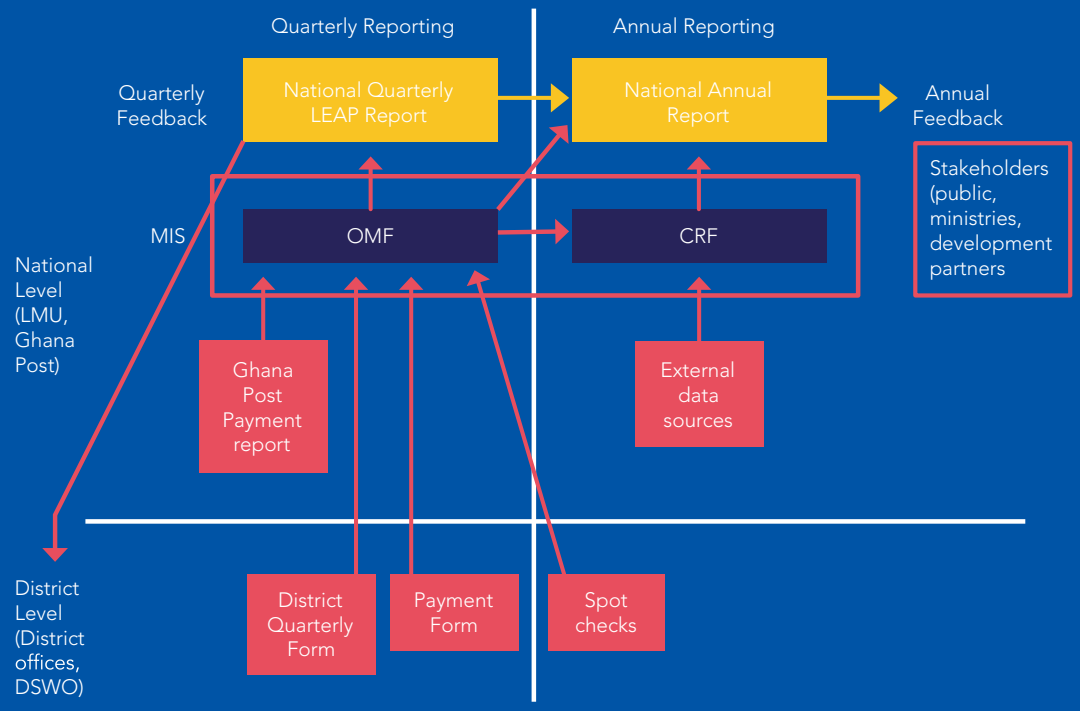
The M&E system has to report on key indicators and provides information and feedback for policy making, planning, budgeting, and implementation. It also has to ensure that the lessons learnt are used to improve the effectiveness of the programme.

Box 20: Continued

Standard reporting formats which respond to stakeholders' information should be institutionalised and used by all M&E actors. Some of these standard formats need to be made accessible to people with disability and made available in Braille, audio, large print, etc. to ensure important stakeholders are not discriminated against based on disability. At the national level, reports are generated quarterly and yearly, in line with the current frequency of report production at the Leap Management Unit (LMU). This should minimise any additional burden on staff. At the district level, quarterly forms and bimonthly payment forms should also be produced, again building on the existing frequency of report production.

Four main reports should be produced and circulated by the M&E system at district and national levels: (i) District Quarterly Forms; (ii) District Payment Forms; (iii) LEAP Quarterly Report; and (iv) LEAP Annual Report. The first two reports are generated at the district level and fed to the national office for the calculation of the indicators that feed the two latter reports produced by the LMU.

Figure 15: LEAP reporting process



The M&E process - Supervision and Feedback

The supervision process puts regional and district officials at the forefront of improvements in LEAP operational processes. It enables operational weaknesses to be addressed swiftly, preventing them from becoming larger operational challenges. However, supervision visits can also be seen as an opportunity to acknowledge good practices on the programme.

Districts should receive regular feedback from the LMU in the form of the LEAP Quarterly Report. This report will be shared with all RSWO and DSWO each quarter.

The LEAP quarterly report, reports on all operational indicators from the District's Payment Forms and the District Quarterly Report Forms. It reveals the operational performance of the district while allowing some comparisons between districts Where urgent and substantive operational issues emerge, the LMU should follow up to help solve this challenge. This process of feedback can therefore be seen as the main mode of communication between the LMU and the districts.

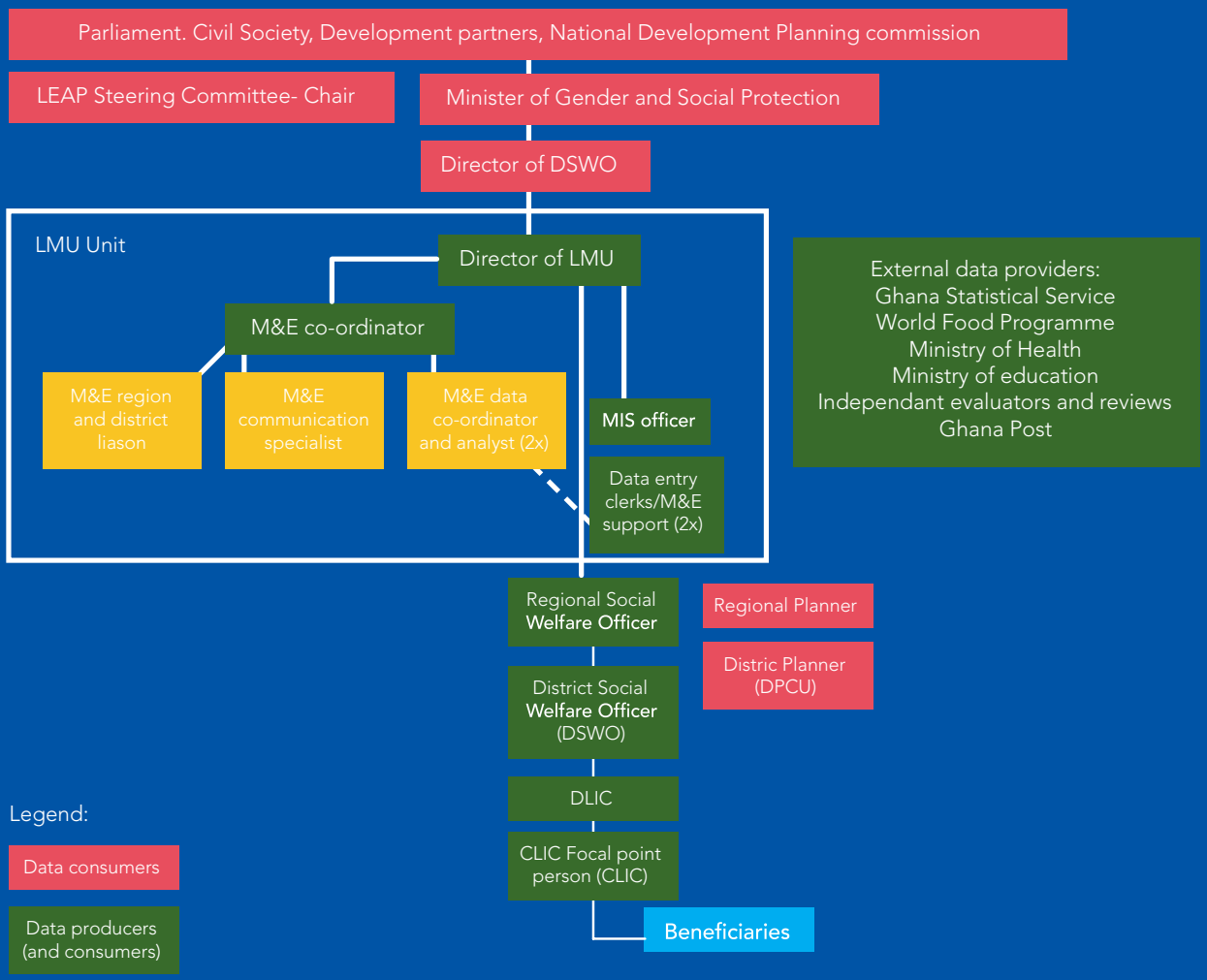
Box 20: Continued

The M&E process - Institutional arrangements

A prior assessment of the M&E situation for the LEAP programme showed how the lack of clearly defined roles, responsibilities and reporting lines led to key M&E functions being side-lined.

Figure 16 below shows the key actors in the proposed LEAP M&E system. Two categories of M&E actors can be identified according to their main role performed: actors who are mainly producers of information and actors who are mainly consumers of information. In addition, four key areas of responsibilities can also be identified in the M&E system: M&E Plan development and management, data collection and analysis, reporting, and dissemination. Actors who are mainly information producers are shown in green. The actors who are pure consumers of data are shown in red.

Figure 16: LEAP M&E actors



Source: Government of Ghana (2013)

2.7 MAKING M&E GENDER-RESPONSIVE AND DISABILITY-INCLUSIVE

As has been noted throughout this chapter, the supply of M&E data needs to be equitable and inclusive of the needs of vulnerable groups. In this section, themes of gender and disability are used to illustrate how M&E systems and activities can respond to vulnerable populations more broadly.

Applying a gender and disability lens to the design and conduct of M&E improves the overall quality of the work as it deepens our knowledge of "what works" for different groups and helps to make M&E activities more context-appropriate, representative and inclusive of diverse perspectives. This inclusion is likely to lead not only to better quality and less biased results but also to greater uptake of the monitored/evaluated programme by recipients. The following five principles are useful for integrating gender and disability in evaluations and can be applied to the inclusion of other vulnerable groups:

1. **Transformative:** Evaluations form part of a larger social protection system and through their conduct and results need to seek to reduce inequalities by gender and disability and support gender-transformative and disability-inclusive change.
2. **Participatory and inclusive:** Diverse stakeholders need to be included in decision-making about what will be evaluated and how the evaluation should be done.
3. **Promoting people with disability, women's and girls' empowerment:** Contextual factors related to gender, disability and power relations, including how evaluations are conducted and by whom need to be considered. Empowerment can be supported by incorporating the insights of diverse groups of people with disability, women and girls in all stages of the evaluation process.
4. **Innovation:** Support innovative approaches to advancing gender equality and disability inclusion and empowerment.
5. **Intentional:** Planning of evaluations need to articulate the purpose and how findings will be used to improve and advance work on gender equality, disability inclusion and the rights of women and girls and people with disability.

By including the voices of women and girls and people with disability in the design and conduct of M&E, the needs of other vulnerable groups are often identified. For example, people with physical disabilities may identify accessibility issues that also apply to older or pregnant people. Young women may identify discrimination issues that are also faced by other vulnerable groups such as ethnic and linguistic minorities. This inclusive lens needs to be applied to the M&E of programmes that specifically target vulnerable populations, those that integrate some equity elements, but also to those that do not focus directly on gender, disability and other issues of vulnerability, because issues of equity are relevant to every programme. This lens must be integrated across all phases of the M&E process. Table 11 outlines some of the practical approaches that can be applied to the planning and conduct of M&E to make data and data collection more gender-responsive and disability-inclusive (UNICEF, 2019). Later on, Table 13 will outline the approaches for later stages in M&E, namely reporting, use and dissemination.

Table 11: Integrating gender and disability in the planning and conduct of M&E

	Gender	Disability
Planning	• Gender analysis integrated into ToC and objectives • Gender equality and rights of women/girls included in evaluation questions • Use of gender equality criterion in evaluations • Gender-sensitive indicators (e.g., gender roles, gender attitudes, time use, agency, decision-making, maternal health, antenatal and postnatal care, menstrual hygiene, etc.) incorporated • Researchers/evaluators with strong knowledge of gender involved; Gender diverse M&E staff are represented on the team • Consultation and collaboration with organisations representing different gender identities • Time and resources allocated for gender-responsive data collection and analysis • Establish recipient feedback mechanisms	• ToC incorporates additional barriers people with disability may face (e.g., accessing cash transfers and information, using cash to improve living standards in line with programme aims) • Measurement of disability is in line with the UNCRPD (e.g., functioning-based questions such as the Washington Group questions rather than binary biomedical questions) • Disability indicators incorporated, starting with SDGs but paying attention to disability issues (e.g., Quality of learning – to overcome the issue of more children with disabilities enrolling in school, but lessons not being inclusive, and learning is limited; the impact of SF on meeting disability-related costs, discrimination, accessibility, etc.) • Researchers/evaluators with knowledge of disability issues involved, including people with disability themselves and their representative groups • The sample size is sufficient to allow for meaningful disaggregation and analysis by disability (and amongst people with disability, e.g., by gender, disability type where possible) • Establish recipient feedback mechanisms
Conduct	• Collection of sex- and gender-disaggregated data across all methods • Gender-responsive data collection (e.g., separate focus groups for women and men) and analysis methods (e.g., disaggregation by gender) are undertaken • Use of mixed methods in data collection • Women, girls, and gender-diverse persons with their organisations participate in data collection and analysis • Uphold ethical and safety standards	• Collection of disability-disaggregated data across all methods • Disability-responsive data collection (e.g., use of disability-accessible formats to allow self-report, training of data collectors on disability) and analysis (e.g., disaggregation by disability) are undertaken • Use of mixed methods in data collection • People with disability and Organisations of People with Disabilities participate in data collection and analysis • Uphold ethical and safety standards

Source: Adapted by authors from UNICEF (2019)

2.7.1 What can happen when M&E is not inclusive?

An evaluation that ignores gender and disability in its design could be perpetuating inequalities and worsening the situation for women and girls and people with disability. For example, evaluations that do not have focus groups conducted separately for women/girls and men/boys when gender-related issues are discussed risk that women/girls will not feel comfortable speaking up or that if they speak up, they might get reprimanded by men in their community who disagree with their views/behaviour. Similarly, evaluations that do not sufficiently consult women and girls in the design stage could end up asking irrelevant questions and produce gender-blind results and consequently, gender-blind social protection programmes. Evaluations also need to facilitate the inclusion of people with disability as respondents for example by offering them the services of an interpreter for those who cannot hear.

Capturing meaningful data on disability can be particularly challenging. Even where data on disability are collected, without careful planning and sampling, the sample of people with disability may be too small to conduct meaningful analyses, particularly where the experiences of persons with different disability types need to be examined. Also, it is important to include village leaders to facilitate access in communities where people living with disability are hidden. Similarly, standard indicators – even SDG ones – may not be sufficiently sensitive to capture disability-specific issues. For example, in a study undertaken in Tanahun, Nepal, household-level data showed that 95% of households with people with disability had safe sanitation facilities (compared to 96% of households without people with disability). However, disability-sensitive analysis at the individual level showed that these facilities were often not accessible and only 60% of people with disability (and 88% of people without disabilities) had safe sanitation access. Gaining this knowledge was only possible because disability-sensitive indicators and analyses were included in the study.

2.8 TAKE-AWAY LESSONS

- Adequate supply of evidence from an M&E system can be achieved when:
- Indicators have been agreed, prioritised and refined as the result of a participatory and iterative process that accounts for the information needs of stakeholders at all levels – including those representing vulnerable populations, as well as reflecting the programme's objectives, Theory of Change, Service Standards and core business processes
- A range of data sources (both internal and external) is adopted, making sure these build on existing sources, minimise the burden of data collection and reporting, and prioritise monitoring over evaluation at the initial stages of programme maturity. Key data sources include: Management Information Systems (MIS); standard periodic reporting from visits, spot-checks, audits, etc; qualitative ad-hoc studies; other administrative databases; Official Statistics data (Census, Household Budget Surveys, Living Standard Measurement Surveys, Labour Force Surveys); externally contracted impact evaluations; Community Monitoring
- Institutional arrangements of the M&E system reflect the overall institutional structure of the programme, work with existing systems, staff and processes, and are built acknowledging the need for time, dedication and in some cases legal frameworks or Memoranda of Understanding
- M&E system and activities apply an equity lens and are inclusive to the needs of vulnerable groups, reducing bias in results and finding out "what works" for different recipient groups.

3

ENSURING DEMAND FOR M&E DATA

3.1 ENHANCING DEMAND FOR M&E

A study conducted by the CLEAR initiative[14] on the demand and supply of M&E information and services in some anglophone African countries[15] concluded that "none of the governments is described as having established a government-wide culture that supports M&E and Performance Management (PM) and the use of M&E and PM findings (…). M&E is often viewed as a control and policing tool or extractive activity, because of how they have been used in the past. This has led to a lack of ownership and little interest in using their findings to inform decision-making. This seems to be most true at local government levels, but also is reflected in line ministries." (CLEAR, 2013)

Unless decision-makers actively seek evidence to support policy-making and programme management, M&E practices are unlikely to take hold. Demand for – and use of - M&E data will be enhanced when:

- At the macro level, the **national policy environment**
  - is 'enabling' (performance-oriented),
  - offers an overall institutional culture that fosters linkages between different stakeholders and has actors focused on planning,
  - allows for Civil Society (e.g., women's groups, Organisations of People with Disability) and Donors to play an active role and participate in the shaping of M&E questions, indicators, methodologies, dissemination and other important decisions.

14 The Clean Initiative
15 The study reviewed M&E practices in Botswana, Ghana, Kenya, Malawi, Namibia, Rwanda, South Africa, Swaziland, Tanzania, Uganda, Zambia, and Zimbabwe.

- At the meso-level, **implementing agencies**
  - have a sufficient level of autonomy in decision-making to ensure M&E activities are perceived as useful and not frustrating (i.e., they are allowed to act on their findings),
  - maintain a strong liaison between central and decentralised levels based on mutual feedback and awareness of location-specific constraints (M&E perceived as learning rather than judgement),
  - backed the process of developing an M&E system in the first place, and have a culture of benchmarking performance across different locations,
  - adopt Standard Service Agreements that help to transparently frame objectives in terms of service delivery (see Section 1 below),
  - understand the potential usefulness of it,
  - do not 'fear' M&E as a 'controlling' function,
  - have sufficient capacity to perform their functions (it is not an added burden to other activities and they have the resources to perform their job).
- At the micro-level, **individuals** responsible for **M&E**
  - understand the potential usefulness of it,
  - do not 'fear' M&E as a 'controlling' function,
  - have sufficient capacity to perform their functions (it is not an added burden to other activities and they have the resources to perform their job),
  - include representatives from vulnerable groups (e.g., women and girls, people with disability, older people, youth, etc.) to make finding more context-appropriate and inclusive.

Best practice internationally includes a balance between 'carrots, sticks and sermons', as summarized in Table 12 below.

Table 12: Incentives for utilisation of M&E - carrots, sticks and sermons

CARROTS	STICKS	SERMONS
• Shift the focus of M&E from 'controlling' to 'learning' • Build forums for local and central level administrators to compare and contrast their experiences (e.g., benchmark across jurisdictions!) • High-level recognition of the goodor best practice • Budgetary incentives forhigh performance • Performance contracts for civil servants & M&E as one criterion for staff recruitment, promotion, and certification • Ensuring that data providers understand how their data are used and the importance of providing accurate and timely data • Training for programme managers and staff	• Enact laws, decrees, or regulationsmandating M&E & formal requirements for the planning, conduct, and reporting of M&E • Withhold part of the funding fromunits that fail to conduct M&E • Achieve greater transparency by regularly publishing information on all programme objectives, outputs, and service quality • Set challenging but realistic performance targets • Involve civil society in M&E of government performance—results in pressures for better performance and accountability	• Use of examples of influential M&E to demonstrate its utility and cost-effectiveness • Frequent repetition of the message of support to and use of M&E at all meetings • Awareness-raising, network building and training on M&E function and its use to deliver better services • Support for M&E from multilateraland bilateral donors in their loansto governments—highlights and endorses M&E

Source: adapted from Mackay (2007)

In **Moldova,** for example, different raions (regions) were compared against each other in the Quarterly Monitoring Reports generated automatically through the programme's MIS, highlighting high-performing and low-performing Social Assistance Departments and allowing to generate a discussion around why that was the case.

In **Mozambique** INAS Delegations (local offices) have a certain degree of autonomy with expenditure and planning with social assistance programmes but budgets are determined based on historical expenditure, rather than performance or needs in the specific geographical area of operation – hampering the usefulness of M&E.

Some examples of strategies to enhance demand for M&E data for decision-making in the social protection sector are reported in the boxes below.

Box 21: Case study - the Monitoring and Evaluation system in SASSA

SASSA operates an organization-wide, centralized Monitoring and Evaluation System, which was developed based on a Stakeholder Information Needs Analysis. This aimed to ensure that only relevant information was being collected. Indicators were selected in a consultative indicator workshop with inputs from the World Bank. The ultimate selection also benefitted from the experience and indicators used by local and international social protection institutions.

The M&E indicators used by SASSA are based on a baseline study covering implementation, administration, and policy management. The number of selected indicators was intentionally kept low, focusing on clear operational objectives and logic for SASSA's operations. All indicators are valid, reliable, sensitive, periodical, and can be disaggregated.

Source: Authors

Box 22: Bolsa Familia Decentralized Management Index

Brazil is a federative republic consisting of the Union, states, municipalities and the Federal District. To correctly implement public policies in the context of this challenging federative arrangement, appropriate strategies are needed to ensure cooperation and coordination between the various actors. The Federal Constitution rules that Social Assistance is a universal social right embodied in those public policies that apply to the entire country and for which different stakeholders share responsibility. It follows that in the Social Assistance sphere all entities are tasked with implementing the relevant policies, including the Bolsa Familia Programme (BFP) and the Unified Registry.

There are two core strategies on which inter-federative coordination of the BFP and the Unified Registry is based: (i) formal commitment to the scheme by federal entities, and (ii) provision of financial support for decentralized management. These mechanisms have enabled the Bolsa Familia Programme to expand systematically over the last ten years in all the municipalities and to ensure that benefits are paid to more than 14 million extremely poor Brazilian families.

The signature of a Term of Adhesion confers BFP membership in Brazil´s 26 states, 5,570 municipalities and the Federal District. This document sets out the standard obligations and responsibilities of each entity that participates in the programme.

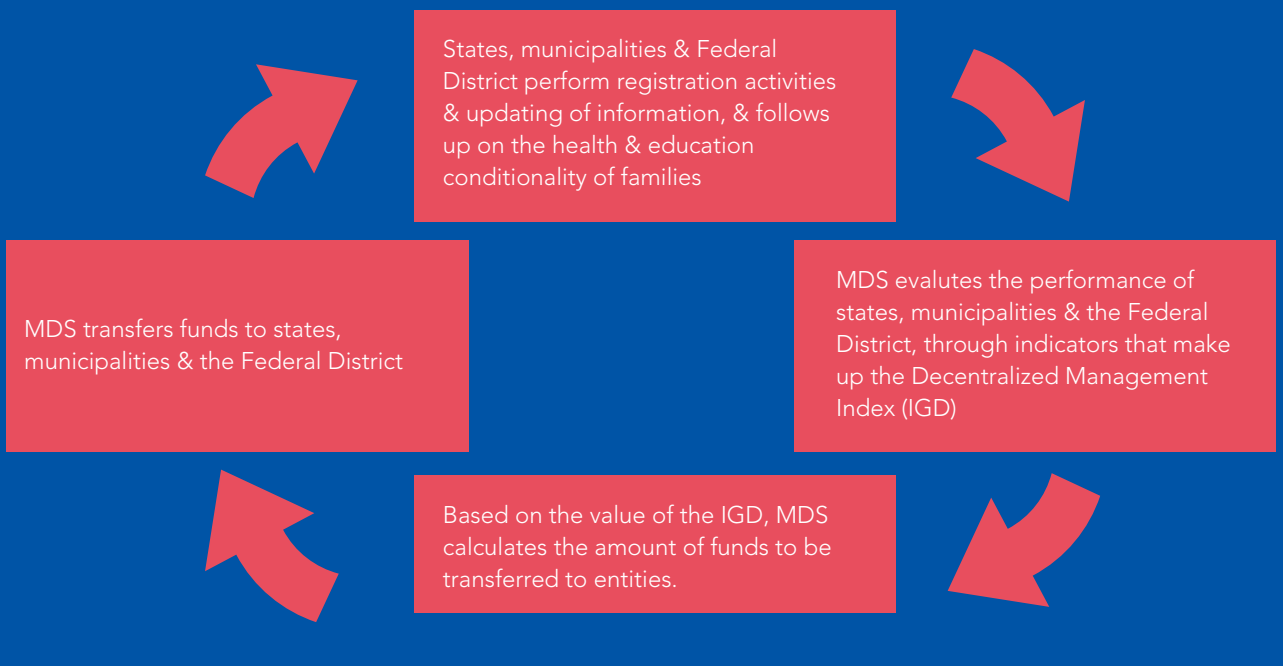
The Decentralized Management Index (IGD) has been adopted by the MDS to support and encourage the federative entities to invest in maintaining and improving the management of the BFP and the Unified Registry. The IGD allows federal government co-financing to be earmarked for states and municipalities, and thus to partially reimburse the costs involved in running the BFP and the Unified Registry. Central government funds feature as revenue in state and municipal budgets and can therefore be directly applied to managing the BFP.

In addition to confirming the obligations entered into under the Term of Adherence, the IGD serves as an indicator for tracking the quality of BFP and Unified Registry decentralized management, as well as a benchmark control for the MDS to release funds to states and municipalities. The higher the value of the IGD, the greater the amount of funds eligible for transfer.

The index serves as a baseline for calculating the value of funds to be transferred directly from the federal government to the municipalities, states and the DF. The following figure shows the IGD as a cooperation strategy for the decentralized management of the BFP.

Box 22: Continued

Figure 17: How the Decentralized Management Index (IGD) works



Source: WWP (2016)

Box 23: The case of CONEVAL in Mexico

CONEVAL is a credible independent agency ensuring transparency in social policy monitoring. In 2001, Congress decided for the first time in Mexican history that all subsidy programes from the federal government would have an annual evaluation. Here, mistrust was a critical aspect. The opposition wanted to prevent the government from using social programmes for political purposes. Toward this end, CONEVAL was instituted as an independent agency, answering to Congress, to monitor poverty and conduct social programme evaluations.

CONEVAL sees the production of evidence as more than just technical work, emphasizing the importance of national dialogue. According to the CONEVAL Director "We started from the basis that if we wanted a poverty indicator to be used by political actors, we could not just request a technical team to deliver a formula and then instruct the politician to 'use this methodology'. In Coneval we say, building a measurement and evaluation system is a political challenge with technical elements, not the other way around. Because if you only take into consideration the technical standpoint – which has to be there - but not the actors, that methodology may be perfect, but nobody is going to use it. There has to be a lot of dialogue".

Source: Dimensions, November 2016, Number 1, Multidimensional Poverty Peer Network (MPPN) accessed, TheDimension Magazine http://socialprotectionet.org/resources/dimensions-magazine-promoting-multidimensional-poverty-measures

Box 24: Using M&E Incentives to Improve Performance – Malaysian Experience

Malaysia implements an approach that focuses on national priority areas and a limited range of outcomes, which include tackling the low income levels through social protection.

Ministerial score cards are developed, tied to the overall planning process in Malaysia and is characterized by a high level commitment from the Prime Minister in the Office of the Prime Minister. In addition to the simple structure for reporting, there is a high level of political support and action. Where there is a low level of performance, Malaysia has introduced an approach to managing the consequences of non-performance called "consequence management". Ministers are ranked according to performance which in practice enhances competition among the different Ministries. Operationally the system is linked to performance meetings between the Prime Minister and Ministers within which scores are mediated. The Prime Minister holds bi-annual Performance Review meetings with Ministers and gives feedback to his Ministers. Annual targets are tied to the 5 years National Development Economic Development Plan and the 10 year Government Transformation Programme.

The system also includes a Ministerial peer review process which is carried out in a systematic manner with Ministers commenting directly on the performance of their peers. As part of the consequence management, there has been instances where Ministers are put in "cold storage (given another post until departure from the service) and others have been removed. Malaysia also has a system of annual reviews of their M & E work and results done by an International Review Panel.

Source: The Clean Iniative - https://www.theclearinitiative.org/sites/clearinitiative/files/20104/african_M&E_workshop.pdf

3.2 ENHANCING THE UPTAKE OF EVIDENCE

Data providers can take explicit efforts to stimulate evidence uptake, by improving:

- **Credibility** can be improved by enhancing the validity, relevance, feasibility, or precision of the M&E information that is generated (see discussion on how this translated in the selection of indicators, data sources and evaluation approaches in Section 2 above)
- **Usability and accessibility** of evidence refer to the tailor-made packaging of the information for the user so that it can be understood and used. If a report meant for the Minister is written in technical jargon, it may not be used and will just be "tossed" aside. Similarly, if a report with relevant information for illiterate stakeholders or people with disability is not produced in user-friendly, accessible formats, such as videos with sign language, info-graphics, Braille, etc., then they will not be able to understand, let alone use the findings. Thus, the data collectors would have wasted their time and resources. It is important to distil the information and tailor the message, medium and communication strategy to different types of target audience.

  Box 25 summarizes key lessons learned on how to improve the usability of results from impact evaluations of cash transfer programmes, based on the experience from evaluations supported by the Transfer Project[16] in eight countries[17] in the southern and Eastern Africa region. Figure 18 further below showcases the power of packaging evidence from M&E by identifying critical policy messages and utilizing effective communication approaches, in this case focusing on challenging false myths and perceptions regarding social assistance programmes.

16 Transfer.Cpc - https://transfer.cpc.unc.edu/
17 Kenya, Ghana, Ethiopia, Zambia, Zimbabwe, Lesotho, Malawi and South Africa

Box 25: The Role of M&E in influencing Social Protection policy-making processes – Lessons learned in Africa

Several factors have been identified as having a positive effect on the influence of the evidence on SCT impacts on policy decisions in the southern and eastern Africa region:

- Evaluations embedded in national policy processes
- Relationship-building and multi-disciplinary research teams
- Messaging and packaging of evidence
- The relationship between demand and supply of evidence (particularly timely availability of key information all along the duration of evaluations)
- The creation of a regional learning agenda, including the establishment of a regional community of practice

While there is considerable diversity across the case study countries, an analysis of the process across the countries suggests that evidence produced by the national evaluations and related research contributed to:

- Building the overall credibility of an emerging social protection sector;
- Strengthening the case for social protection as an investment, not as a cost, and addressing public perceptions and misconceptions;
- Supporting learning around programme design and implementation to inform programme improvement in key areas such as targeting, access, transfer size, and the role of complementary activities; and
- Shaping policy discussions beyond the national context and informing regional social protection agendas.

Source: Davies et al. (2016)

- **Access:** In some cases, the data users may want to use evidence for decision-making but don't know where to find the information. M&E data must be readily available to the different users. The information may be uploaded on a website or portal so that both internal and external users can use it to make decisions. The transparency ensured by Kenya's online data from the Single Registry is a good example of improving access to M&E data (see Box 26). Also, providing local governments with written copies of findings from analysis of M&E data may be helpful in some African contexts where electricity or the availability of computers may be a challenge.

Box 26: Enhancing Access to Information from the Single Registry of Social Assistance Programmes in Kenya

The Single Registry of Social Protection Programmes in Kenya is a software platform designed to manage and provide integrated oversight of the principal social assistance cash transfer programmes in Kenya. The Single Registry has evolved as part of the broader social protection policy framework.

The Single Registry consolidates data from programme management information systems (MIS) of Kenya's largest social cash transfer programmes, into a unified recipient MIS that aggregates and analyses this information to support planning, coordination, accountability and both programme and policy-level decision-making.

The data will be most used by the local and national government, policymakers, safety net programme designers and managers. The Single Registry also provides an accessible reporting system and easy-to-use dashboards that generate reports on many aspects of the programmes' performances. It can show if recipient needs are being met, and how targeting of recipients and cash transfers might be better harmonized and aligned and overlap reduced. This information is critical for the effective management of the cash transfer programmes and for the continued development of social assistance policy.

Box 26: Continued

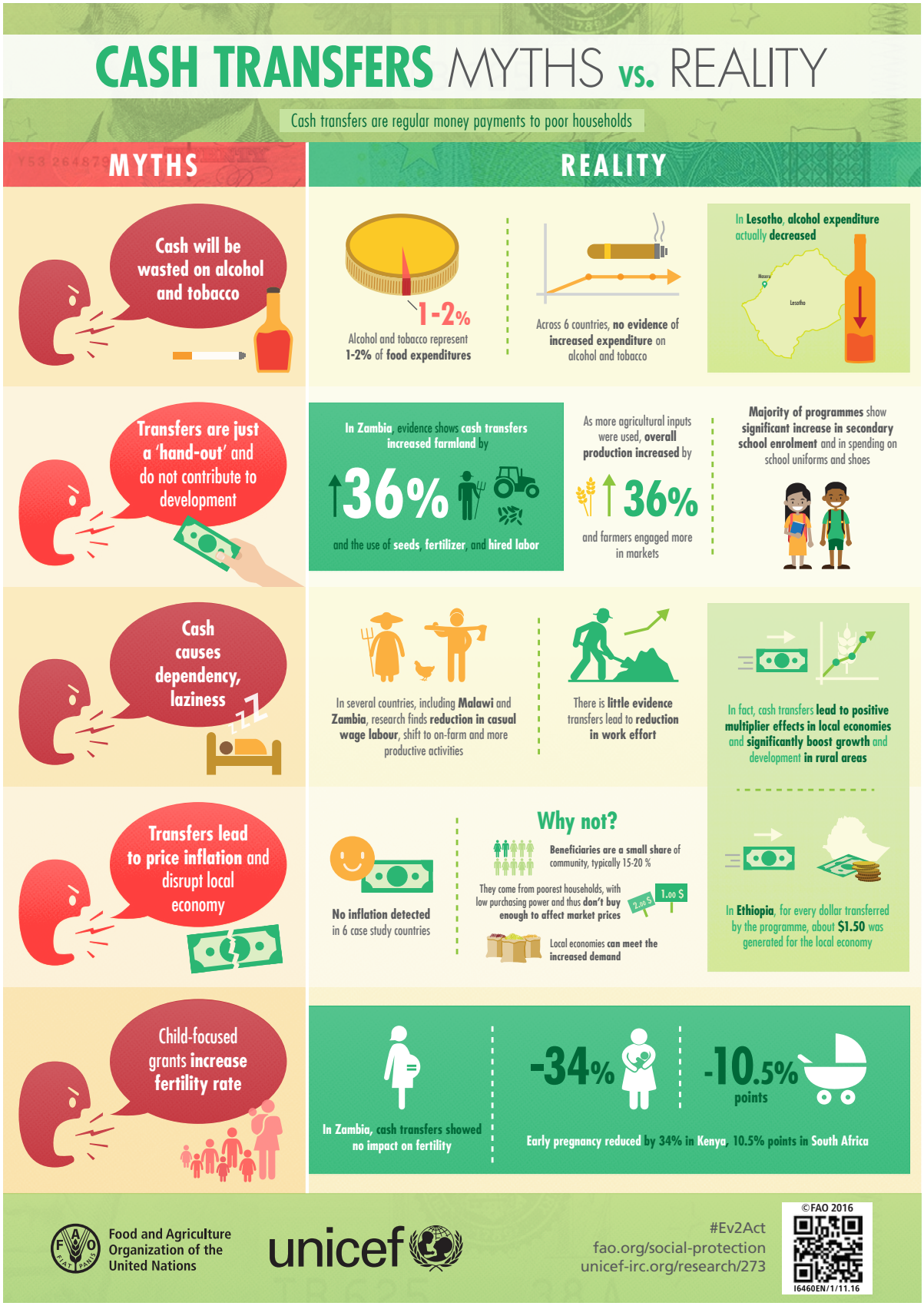
Anyone can use the Single Registry as a tool for searching data on Kenya's social protection programmes and beyond. Data from the Kenya Single Registry, including summary statistics, information on recipients, payments, the complaint systems, maps and reports can be accessed online at socialprotection.go.

The introduction of the Single Registry has increased the transparency, accountability and functioning of the social assistance programmes in Kenya. Previously, managers were unable to answer the simple question of who was receiving what within each social protection programme. The registry has now linked data from disparate programmes into one coherent viewing platform, providing a means of managing and cross-referencing the information.

Similarly, information from the Ghana LEAP cash-transfer programme M&E dashboard is available online (read more in Box 5 above).

Source: http://mis.socialprotection.go.ke:20301/

Figure 18: Infographic about myths and misconceptions on social cash transfers in Africa



Source: Transfer Project, transfer.cpc - https://transfer.cpc.unc.edu/

3.2.1 Enhancing the uptake of gender-sensitive and disability-inclusive evidence

The credibility, usability and accessibility of evidence will be increased when it incorporates the perspectives and needs of vulnerable groups, including women and girls and people with disability. Table 13 outlines best practice approaches to integrating gender and disability in the use and dissemination and reporting of M&E findings. Such approaches are becoming increasingly expected not only by civil society but also by donors and other partners involved in the design and financing of social protection programmes.

Table 13: Integrating gender and disability in the use, dissemination and reporting of M&E findings

	Gender	Disability
Use and dissemination	• M&E results and recommendations reach all gender identity groups and key implementing partners • Dissemination methods and formats are gender-sensitive (e.g. avoid negative gender stereotypes) and use user-friendly formats (video, photos, social media) to reach recipients and their organizations • Dissemination conducted in a gender-responsive and culturally-appropriate manner • Gender-responsive recommendations are incorporated to improve programme delivery	• M&E results and recommendations reach persons with diverse disabilities and their representative organizations • Dissemination methods and formats are disability-inclusive (e.g. Braille, audio, large text, video with sign language, simple language, etc.). • Collaboration with Organisations of Persons with Disabilities, other groups and local governments for dissemination and uptake • Disability-inclusive recommendations are incorporated into programme delivery • Using M&E results to track progress against SDGs and the UNCRPD
Reporting	• Reporting on gender equality issues included in all reports, including the recommendations • Unexpected results/outcomes on gender equality are addressed	• Reporting on disability issues included in all reports, including the recommendations • Unexpected results/outcomes on disability inclusiveness are addressed

Source: Adapted from UNICEF (2019)

3.3 EXPERIENTIAL LEARNING AND THE LEARNING ORGANISATION[18]

Under the traditional M&E approach most projects, including in the social protection field follow a standard practice as articulated in project planning or logical framework approaches and define a 'development project' as inputs (financial and other resources), which are translated by an implementing agency into specified activities to produce useful outputs. These outputs have the goal of outcomes and impacts of higher well-being for the intended recipients.

In the last ten years, there has been an accelerating rise in the criticism of traditional M&E and a corresponding rise in the prominence given to the use of rigorous techniques for project evaluation. The criticism of M&E practice has two key elements: a) evaluation was too ex-ante and needed to be more ex-post, and b) evaluation should be more focused on the impact on outcomes, not just inputs, and based on a rigorous counter-factual. This has led to a significant rise in focus on rigorous impact evaluations (see more in Box 12 above).

There are three fundamental reasons why both the traditional M&E approach and the more recent impact evaluation approach fall short of the learning needs of most organizations:

- **High-dimensional and complex design space implies that learning** 'what works' has to be flexible and dynamic

18 This Section is based on material from Pritchett (2013) and the PDIA initiative available at Harvard.edu/

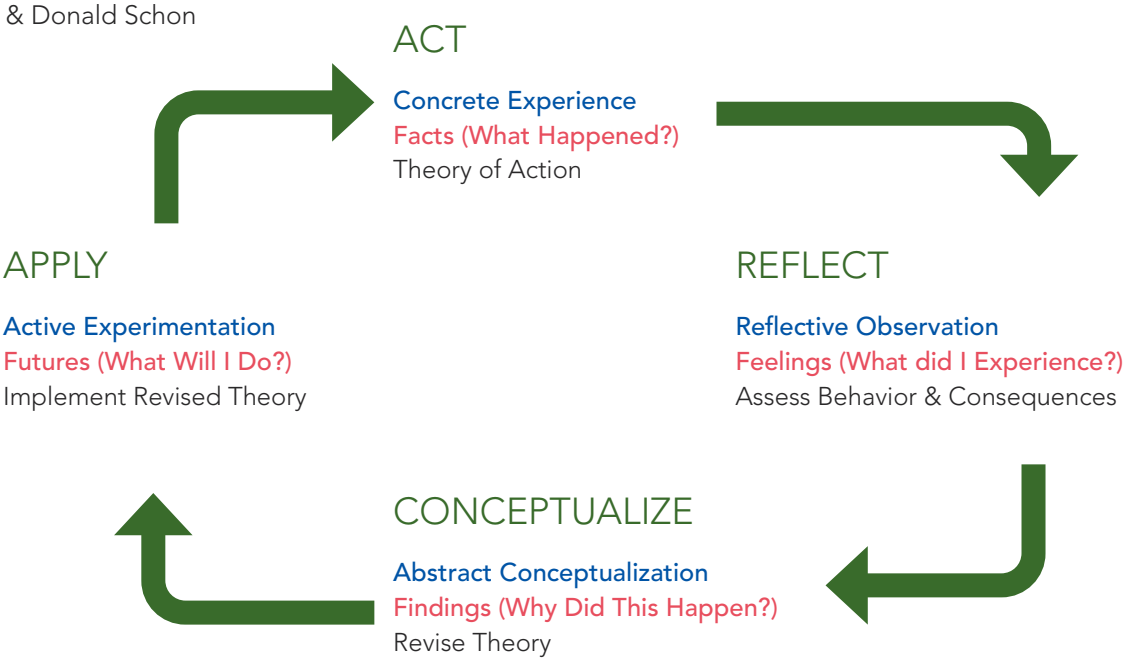
- Many development **problems are problems of implementation**—moving from inputs to outputs (for which an impact evaluation that measures outputs to recipients is not yet needed)
- Like human beings, organizations and systems **learn through experience**, and not (only) through evidence

3.3.1 Learning through Experience

People learn through a circular process of action, conceptualization and evaluation. It involves referring to previous experiences as well as anticipating outcomes. What we do is the result of observation, action, and reflection. Our behaviour reflects how we compare experiences from the past, deeming them good or bad, successful or unsuccessful. We also look at others - what actions of theirs are good? And last but not least, we develop unique strategies that apply best to our specific situation. Experiential learning is the process of experience and is more specifically defined as "learning through reflection on doing" (Kolb 2014).

Figure 19: Experiential Learning Cycles

1. David Kolb
2. Roger Greenway
3. Chris Argyris & Donald Schon



Source:ed batista - https://www.edbatista.com/2007/10/experiential.html

Development practitioners are well aware that a lot of learning from a project happens after the design, but well before any formal 'evaluation' but as it is this learning is often haphazard and below the radar. The participation of different stakeholder groups is integral to providing insights about the appropriateness of the methodology and interpretation of findings to enable continuous learning and programme improvement. This is particularly relevant for vulnerable groups whose needs are often not captured appropriately or at all. Engaging with women's groups, youth, Organizations of Persons with Disabilities and other civil society representatives not only ensures that their priorities are captured in data collection but also creates early buy-in, establishes the legitimacy of the M&E activity and team, increases the uptake of evidence and can reduce criticism that the evaluation (and programme) is gender-blind or not inclusive of disability.

The goal is to bring the currently informal processes of experiential learning, from project implementation, explicitly into the overall strategy of development organizations. Prittchett et al. (2013) propose to explicitly add a new 'e' in MeE, defined as structured experiential learning. This is the process through which an organization learns during the period of project implementation.

3.3.2 Towards structured experiential learning

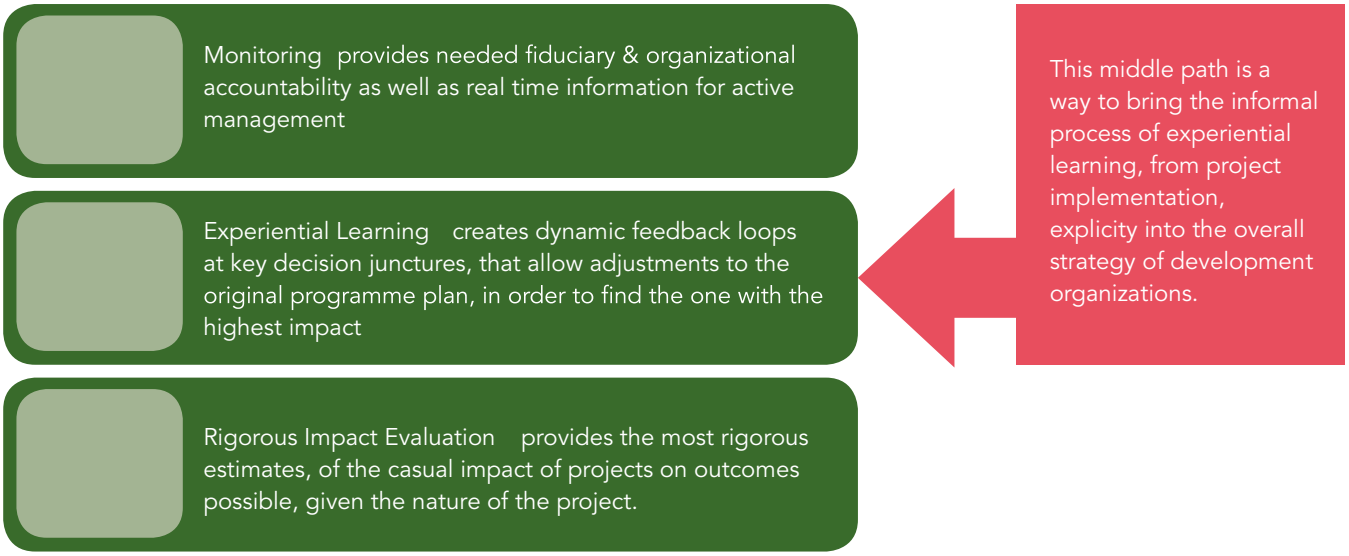
Many implementing agencies (or at least significant proportions of the people in those agencies) want to do what they want to do and do it well if possible. The difficulty is that the idea of 'independent evaluation' often arises when a principal (e.g. funding agency) wants to select among alternatives and provide more support over time to 'what works'. When organizations (correctly) perceive that the role of the evaluator is to be an instrument to cut their budget if they are 'ineffective' rather than help them be effective, the enthusiasm for evaluation naturally wanes (Pritchett 2002). Therefore, implementing agencies often are less than enthusiastic (even subversive) about rigorous impact evaluations of outcomes.

However, implementing agencies are often interested in the evaluation of what works to produce outputs. The management of an implementing agency has some control over the outputs of a development project by managing inputs and activities. Organizations would allow evaluations which focused on outputs because outcome data is costlier than output data, it nearly always involves engagement with actors who are external to the development project.

Structured experiential learning is the process of disaggregating and analyzing data on inputs, activities and outputs chosen to be collected by the project to draw intermediate lessons that can then be fed back into project design during the project cycle. The idea is to take the key insight about using randomization and other rigorous methods to identify the impact and expand it dramatically—at a lower cost—by using the development project itself as a learning device. Variations in alternatives within the design space within the project can be used to identify efficacy differentials in the efficacy of the project on the process of inputs to outputs, which can be measured at low incremental cost at high-frequency intervals, for real-time feedback into implementation, at key decision junctures. Rather than thinking of projects as a single element of the design space, projects that are intended to be innovative are authorized strategic evidence-responsive crawls over (part of the) design space.

To maximise experiential learning it must be anchored on an organizational learning strategy which consists of a project-specific mix of Monitoring, Experiential Learning and Evaluation (M&E) (see Figure 20 below).

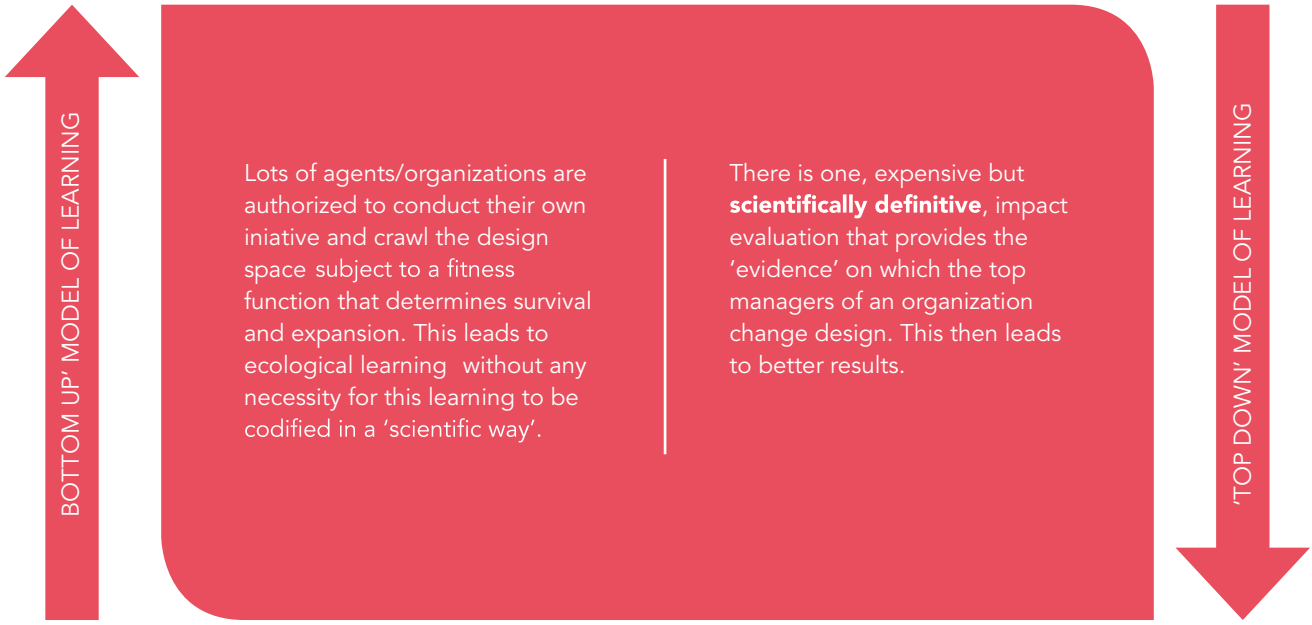
Figure 20: The role of Experiential learning in M&E



Source: Prichett et al (2013)

Often the implicit model behind an impact evaluation approach is that the management of the implementing organization will change the design based on **entirely technocratic 'evidence'** and that implementation will change by edict from above. An alternative is that implementing agents will change their behaviour when they are convinced that the new behaviour furthers their objectives, which include both self-interest but also some concern for the organization's outputs and outcomes. If this is the case then the involvement of the implementing agency and agents in the learning process is essential to the impact of the learning (see Figure 21).

Figure 21: Top-down versus bottom-up learning



Source: adapted from Prichett et al (2013)

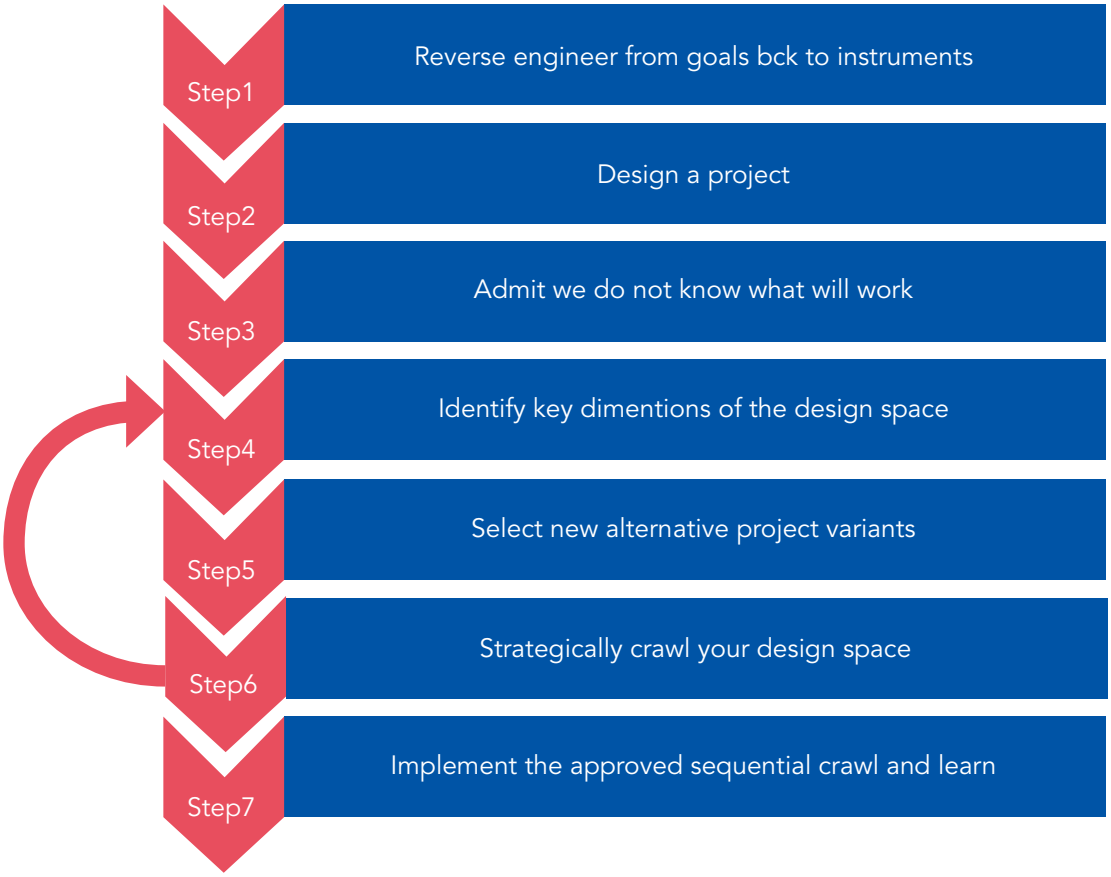
The reality is that with complex endeavours no one can know what will work in advance. Development project managers do not know if the inputs will lead to useful outputs (internal area within their control) or if the outputs created will in turn lead to outcomes and impacts (not within their control). As we have argued above, given the level of granularity at which projects have to be designed cannot be 'evidence based'— even if one draws on all of the available information. Development projects are not like chemistry—which is complicated but not complex—where we can predict exactly how interactions will work under specified conditions because we have empirically validated invariance laws that cover all the relevant contingencies.

Some projects are just logistics, the solutions have been tried out and proven in context (both overall and organizational), and hence the purpose of the project is just scaling. However, not all projects are just the logistics of implementing known solutions and hence processes that insist that all projects present themselves either as logistics or as small-scale pilots or field experiments create unnecessary friction and confusion.

Dealing with complexity requires a different approach to programming as well as monitoring and evaluation, challenging the traditional wisdom that change happens linearly. It requires embedding in the process opportunities for iteration, feedback, incorporating the perspectives of vulnerable groups and continuous learning (see Figure 22). From an M&E perspective dealing with the complexity required to overcome the rigidities of the traditional log frame approach and adopt a more flexible framework for searching and assessing solutions based on continuous practice (see Box 22).

Read more about 'Doing Problem Driven Work' and Problem Driven Iterative Adaptation here: Harvard.edu - https://bsc.cid.harvard.edu/

Figure 22: How to go about experiential learning?



Box 27: SearchFrames for Adaptive Work (More Logical than Logframes)

Although the benefits of experimental iteration and learning process seem very apparent to most people, many development organizations make it difficult for staff to pursue such approaches, given the rigidity of log frame and other linear planning methods. Funding organizations demand the structured, perceived certainty of a log frame-type device and will not allow projects to be too adaptive. In response to this concern, we propose a new log frame-type mechanism that embeds experimental iteration into a structured approach to make policy or reform decisions in the face of complex challenges called the SearchFrame.

The SearchFrame facilitates a transition from problem analysis into a structured process of finding and fitting solutions. An aspirational goal is included as the end point of the intervention, where one would record details of 'what the problem looks like solved'. Beyond this, key intervening focal points are also included, based on the deconstruction and sequencing analyses of the problem. These focal points reflect what the reform or policy intervention aims to achieve at different points along the path towards solving the overall problem. More detail will be provided for the early focal points, given that we know with some certainty what we need and how we expect to get there. These are the focal points driving the action steps in early iterations, and they need to be set in a defined and meaningful manner (as they shape accountability for action). The other focal points will reflect what we assume or expect or hope will follow. These will not be rigid, given that there are many more underlying assumptions, but they will provide a directionality in the policymaking and reform process that gives funders and authorizers a clear view of the intentional direction of the work.

The SearchFrame does not specify every action step that will be taken, as a typical log frame would. Instead, it schedules a prospective number of iterations between focal points (which one could also relate to a certain period of time). Funders and authorizers are thus informed that the work will involve a minimum number of iterations in a specific period. Only the first iteration is detailed, with specific action steps and a specific check-in date.

Box 27: Continued

Readers should note that this reflection, learning and adaptation make the SearchFrame a dynamic tool. It is a tool to use on the journey, as one makes the map from origin to destination. It allows structured reflections on that journey, and report-backs, where all involved get to grow their know-how as they progress, and turn the unknowns into knowns. This tool fosters a structured iterative process that is both well suited to addressing complex problems and meeting the structural needs of formal project processes. As presented, it is extremely information and learning intensive, requiring constant feedback as well as mechanisms to digest feedback and foster adaptation on the basis of such. This is partly because we believe that active discourse and engagement are vital in a complex change processes, and must therefore be facilitated through the iterations.

Source: Andrews et al. (2017), see also building state capability - https://buildingstatecapability.com/2016/06/06/searchframes-for-adaptive-work-more- logical-than-log frames/

3.4 TAKE-AWAY LESSONS

- M&E systems must be designed in such a way to strike a balance between the capacity to produce quality evidence in a timely fashion and the demand for evidence of a particular kind needed for decision-making by multiple users.
- To maintain and sustain the supply and demand balance, there should be constant dialogue between data providers and data users coupled with ensuring that the evidence is made usable and accessible to key stakeholders for policy making and the community and providing incentives to the data users to stimulate the uptake of evidence.
- On the demand side of M&E, it is important to create a culture of learning and not blame and this way, it ensures the usefulness of the M&E framework to its key users.
- A good M&E system is critical to safeguarding compliance with existing legislation, ensuring transparency and accountability and building a basis for the continuous improvement of social protection systems. A good M&E system promotes a continuous learning cycle, fosters transformation in social protection, and improves service delivery.
- To maximise learning during implementation, organisations should use structured experiential learning and adopt a participatory approach with key stakeholders, including vulnerable groups. In the face of complex challenges, change is generally not a linear process, requiring experimental iteration and frequent feedback loops.

4

SERVICE STANDARDS, PERFORMANCE MANAGEMENT AND STRATEGIC PLANNING

4.1 SERVICE STANDARDS AND SERVICE CHARTERS

Service Standards outline the specific delivery targets established by an organisation, and are made up of a set of commitments that an organization promises to honour when delivering a service. They also describe what a client or user can expect to receive from the service and how the service will be delivered. A well-designed **performance management system** associated with those Standards (and adequately monitored by the M&E system, see Section 1) is essential to ensure a cost-effective and high-quality service that meets the needs of service users. Specifically, Service Standards aim to:

- support the provision of consistently high-quality service delivery
- encourage continuous improvement and identify specific areas for improving service quality
- assist service providers to self-audit the quality of their service
- foster a collective commitment to quality through a common set of clear and measurable criteria
- assist users in knowing what to expect from service providers concerning the quality of service delivery
- maximise staff satisfaction and confidence with the service
- meet reporting, transparency and accountability requirements
- assist with monitoring and evaluation processes.

Best practice in the establishment of standards in this sector include the following:

- Tailoring the standards to their focus. As useful guidance in the social assistance sector, see the World Bank's (2011) suggested "governance and service quality" standards and related indicators for Safety Net Programmes, focusing on human resources, financing and resource management, operational procedures, organizational performance and overall quality of delivery.
- Ensure standards guarantee (World Bank, 2011):
    - *Service accessibility:* the extent to which social assistance services are physically, informationally, spatially and socio-culturally accessible to all individuals in the target population, regardless of their age, gender, ethnicity, language, disability condition, location, and so on (See Table 12 below for some examples of rough indicators for this).
    - *Privacy, dignity and confidentiality:* the extent to which each recipient is treated with the same level of privacy, dignity and confidentiality.
    - *Cultural awareness:* delivery of non-discriminatory services which are sensitive to the social and cultural values of the individuals, their families and their community.

- *Complaints and appeals:* the existence of publicized and easy-to-use complaints/grievance procedures, including a commitment to deal with and solve complaints within a time limit. See MODULE ADM.
- *Exit and re-entry:* extent to which individuals are assisted to plan for exit from assistance and assured that re-entry is available if required.
- *Service integration:* development of links with other social assistance service providers at local, state and national levels to ensure access to complementary services.

Table 12: Example of potential dimensions of service accessibility to monitor

DIMENSION	SUB-DIMENSIONS AND ROUGH INDICATORS
Service Accessibility	**Accessibility Policies** • Existence and application of equal opportunities and disability policies, procedures and training forservice delivery • The extent to which the target population is defined, their needs regularly identified and services are planned and delivered to meet those needs **Affordability** • Availability and use of different means to reduce the cost of programme information and application (e.g., mobile offices, toll-free phone numbers, accessible formats for people with disability, etc.) • Travel time and transportation costs to facilities, including for people with disability (e.g., due to inaccessible public transportation) • Application and enrolment costs, including caregiver costs (for people who need assistance to enrol) • Average time required for programme application and enrolment • Individual's perception of ease of service accessibility and paperwork required **Physical, spatial and functional accessibility** • The convenience of location and access (% of facilities close to public transport, with clear signage, level access, that are compliant with accessibility regulations for people with physical disabilities; % of services that can be accessed from home/community) • The convenience of facilities (e.g., % of facilities with adequate waiting areas which accommodate women, children, the elderly, and people with disability, are properly maintained, sanitary, completely operational, fully stocked and supplied) • The extent of use of technology to provide information and access to services (if appropriate) • Posting and observance of operation hours (e.g., % of facilities with posted operation hours, % of facilities with staffed information/registration counters during business hours at random visits) • The convenience of operation hours (% of recipients who consider operation hours convenient) • Availability of information in disability-accessible formats, including Braille, large text, simple language and info-graphics, audio, video with sign language, etc. **Socio-cultural accessibility** • Availability of information materials in different languages and formats, including pictures and info-graphic for people with limited literacy • The ratio of the number of recipients who have difficulties communicating in the official language andthe number of trained interpreters or bilingual staff • Availability of information and tailored service access for individuals with special needs (e.g., people who have difficulties understanding or reading the official language, people with limited literacy or learning difficulties) • Individual's perception of service provision in a manner sensitive to the age, sex, ethnic, and linguistic background of each person

Source: World Bank (2011)

- *Ensuring standards measure* what matters to service users (through research and consultations).[19]
- *Involving staff at all levels in determining the standards and their related measures*. Measures that are seen by staff as irrelevant, unrealistic, inappropriate or unfair will be counterproductive. For example, they may focus on the measure given at the expense of other more important factors, they may try to get around the system, or they may sub-optimise or concentrate on short-term issues.
- Making sure there is a *balanced assessment and set of measures,* to avoid managers focusing all of their attention on one aspect of service delivery, rather than on the bigger picture.
- Developing performance measures and related indicators based on the following principles:
  - Including both perception measures and ('objective') performance indicators.[20]
  - Using a combination of outcome and process measures.[21]
  - Ensuring benefits outweigh the costs of obtaining/calculating them.
  - Developing clear systems for translating feedback into a strategy for action that can be communicated to staff.

Service Standards can also be translated into a 'Service Charter' that includes and defines the key values of the administration (for instance: citizen friendliness and service-oriented attitude, sensitivity, humane attitudes, efficiency, transparency and integrity), while also defining citizen's rights and responsibilities and key performance targets in terms of service delivery (see also Box 28).[22]

19 See also Moullin, 2004; Osborne, Radnor and Nasi, 2013

20 Perception measures are obtained directly from service users and other stakeholders, while performance indicators are recorded directly by the organisation. For example, measuring the average time patients wait in a GP's surgery – a performance indicator – is important, as this will show whether the actual waiting time has improved. However, this will not tell the surgery how satisfied or irritated, a particular patient is with the length of wait. A carefully designed patient questionnaire or a focus group (examples of perception measures) would indicate this, and therefore both types of measure are needed. Another advantage of perception measures is that they can pinpoint changing expectations.

21 Measuring outcomes (e.g. whether an ill patient recovers) is important because they are of vital importance to patients and service users. Similarly, process measures are important because they measure the way service is delivered, which also matters to patients and service users. One problem with outcome measures is that they may only be available for several months or years following a particular treatment, by which time the personnel or the treatment regime may have changed. They also cannot be used to detect near-misses. However, there is also a danger in using process measures if these are not linked to outcome measures or patient/user satisfaction. A service may then conform to process measures used, but bear little relation to patient outcomes or satisfaction.

22 international budget

Box 28: Case Studies - Service Charters

In Namibia, a Service Charter is made available as a pamphlet and includes telephone numbers and contact details for a 'complaints coordinator' at the relevant ministry. It also includes postal addresses and phone numbers for regional and local offices and sub-divisional heads in the parent ministry.[23] The Namibian Service Charter clearly stipulates the processing times for certain activities.

Table 13. Extract from the Namibian Ministry of Labor and Social Welfare's Service Charter[24]

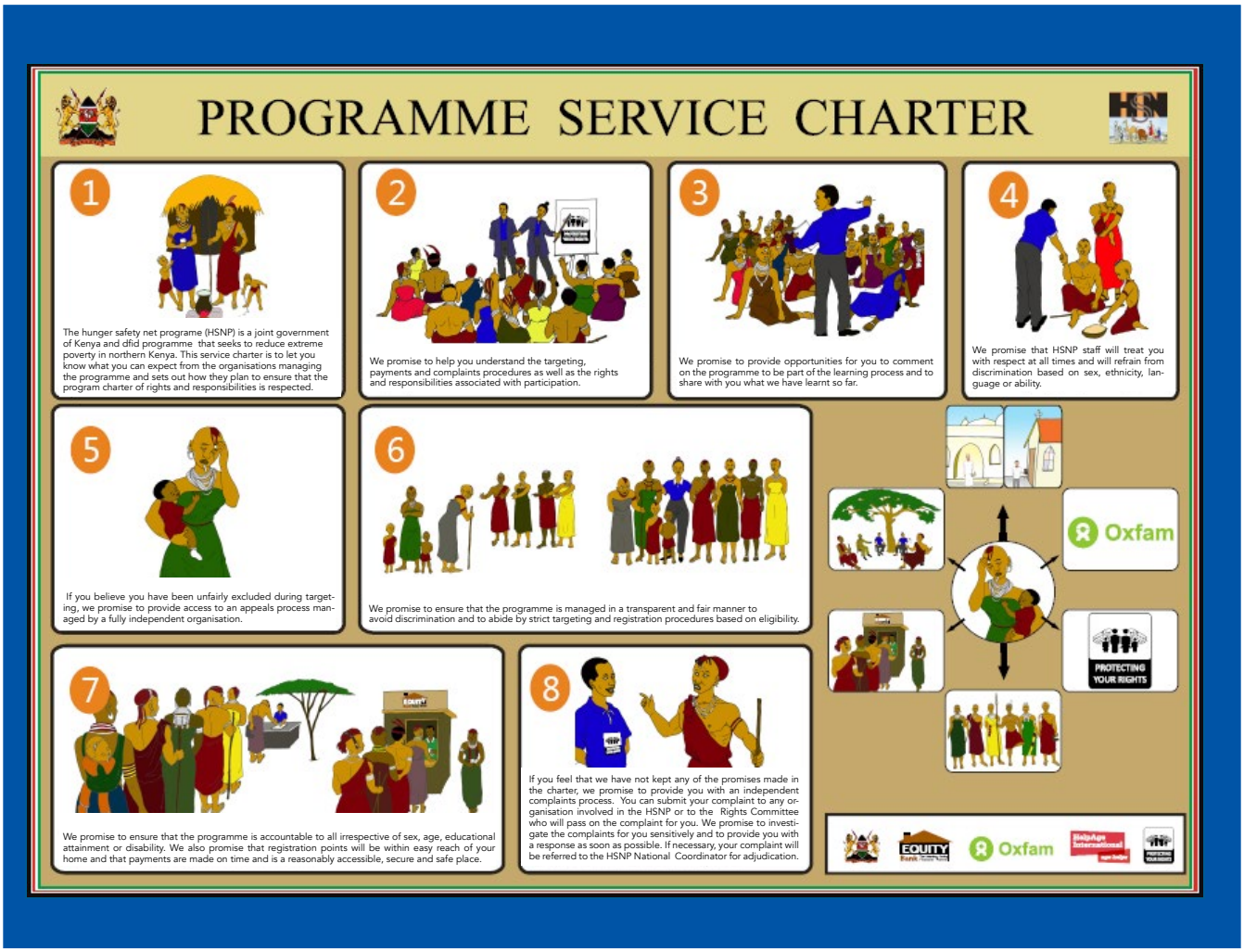
ACTIVITY	PROCESSING TIME
Payment of Old Age Disability grants attestation with	90 days
Changes in pay points and modes	60 days
Effected address changes	24 days
Finalise reinstatement	90 days
Payment of insurance claims	48 days
Handle enquiries	3 days
Deal with queries	14 days

In South Africa, delivery of public services follows the principles of 'People First', which encompasses nine core values for guidance of public servants: consultation with citizens; setting service standards; increasing access to information; ensuring courtesy; providing information; openness and transparency; redress; and value for money. In an effort to contribute to service improvement, People First requires departments to benchmark the standard of service delivered to citizens. Service standards are measured by public satisfaction inquiries regarding services delivered.

In Kenya, the HSNP's programme Service Charter was widely communicated using a simple visual support (see image below).

23 ILO, 2014, p. 159
24 p. 160

Box 28: Continued



Source: Namibian Ministry of Labor and Social Welfare Service Charter, HSNP Service Charter, South Africa 'People's First principle's

4.2 PERFORMANCE MANAGEMENT AND STRATEGIC PLANNING

Implementing Service Standards implies linking these to Strategic Plans and Performance Management: which help to ensure that service goals are consistently being met effectively and efficiently. This is not an easy task, often resulting in resistance on behalf of those who are being 'evaluated'. Box 29 explores some of the reasons why this is the case.

Box 29: Resistance to performance management linked to service standards

Programme managers often fear that elected officials, interest groups, and the media may use service quality and programme outcome information as fodder for attacks on them. They are "concerned that such information stands a good chance of being misused and that they will be blamed for any negative findings" (Hatry, 1992).

This resistance can translate into several types of behaviour including (Moullin, 2009; Brooks, 2007; Hechman et al, 2002):

Tunnel vision: for example, if the standards focus on Outcome X, staff will have less incentives to pursue Outcomes Y and Z – even though they may be just as important.

Gaming: this implies changing behaviour to cheat the system and not to improve performance. For example, when the NHS in England introduced waiting time targets in emergency departments, patients were made to wait in ambulances.

Data misrepresentation: actively (and misleadingly) changing the data that is shared with higher levels to comply with standards.

Complacency or 'convergence': for example, the assumption that standards set the boundaries of best practice and not striving to perform better. In the case of social assistance, this can mean a 25-day wait for application processing to be deemed 'acceptable' because it is within the 30-day range set by the standards.

Cream-skimming: selecting recipients/clients/users based on the likelihood with which these will help to reach targets rather than their actual needs. For example, if a target involves bringing X% of people over the poverty line it will be easier to do so by cream-skimming those who are relatively richer.

Sub-optimisation: the pursuit by managers of their narrow objectives at the expense of strategic coordination.

'**Ossification':** a disinclination to experiment with new and innovative methods. For example, shallow and narrow teaching in pursuit of gains in test results, or cutting funding for R&D.

Myopia: i.e. not focusing on the long term (this is a variation on the 'tunnel vision' above). For example, if targets for new applicants for a benefit need to be met there will be a tendency to disregard people who are applying a second time (making them wait).

Tokenism: e.g., paying lip service to issues like gender equality and disability inclusion, but doing little to change the situation and experiences of members of these groups.

These attempts to force the system are not necessarily the fruit of bad will on behalf of the staff involved. In some cases, these attitudes are due to performance standards that are applied inflexibly without taking into account external factors such as local conditions and challenges (i.e. things that are not within programme control) (Hatry, 1992).

According to the literature, **best practice** implementing service standards dictates the following:

- Apply performance standards consistently, but with some flexibility - allowing for managers to provide explanatory information with the programme performance data.
- Involve as many staff as possible in the process of developing and refining the service standards and related performance system (building ownership).
- Develop a relevant, useful and timely reporting system (see Section 1 on the M&E system):
    - Making sure performance measurement against standards is reported regularly (and at useful times for the manager: i.e. when a decision needs to be taken);

- Making sure data is presented in an adequately disaggregated way to reflect managers' needs at that level and to dig deeper into how outcomes are linked to demographic and service characteristics in a given area;
- Allowing for the comparison of performance across different units and across time to generate benchmarking and 'yardstick competition';
- Adding depth of interpretation to the information by linking it to additional qualitative research or efforts to identify more specifically where the strengths and weaknesses of the services are located;
- Reporting the data clearly and interestingly (bar graphs can work better than tables as they have a visual impact, etc.); and
- Continuously training managers and staff to interpret and use that data (this cannot be a one-off exercise).

- Focus the measurement system on continuous improvement for service users, not on a 'blame culture'. If performance on a particular measure is below the bar, the emphasis needs to be on establishing what went wrong and how this issue can be addressed in the future.
- Link incentives to performance evaluation and management (bonuses, etc.)
- Collaborate with representative stakeholders and groups (e.g. women's groups, youth, Organizations of Persons with Disabilities by having them on advisory committees and boards, as part of the audit and M&E teams, etc.) to ensure their perspectives and needs are incorporated and the service is accountable to them.

4.3 TAKE-AWAY LESSONS

- Service Standards outline the specific delivery targets established by an organization, and are made up of a
- set of commitments that an organization promises to honour when delivering a service. They also describe what a client or user can expect to receive from the service, and the manner in which the service will be delivered.
- In designing these it is essential to ensure standards measure what matters to service users (through research and consultations) and involve staff at all levels in determining the standards and their related measures.
- Service Standards can be also translated into a 'Service Charter' to be widely circulated. Most importantly they should be linked to Strategic Plans and Performance Management, helping to ensure that service goals are consistently being met in an effective and efficient manner.

5

ACCOUNTABILITY MECHANISMS

In MODULE ADM we discussed the core accountability function played by a programme-specific or sector-wide Complaint and Appeal Mechanism for social protection delivery. A wide variety of further internal and external measures can be set up to increase the overall accountability of service provision in the social protection sector, in combination with a strong M&E system (described in Section 1 above). Two key objectives are served by these measures: ensuring and improving quality service delivery as well as enhancing the impact of social protection provision. We discuss the most important below.

5.1 INTERNAL MEASURES

5.1.1 Spot-checks, Supervisions and Inspections

Beyond the complaints and appeals mechanism, other internal procedures can be set up to ensure compliance with standard policies, legislation, processes, service standards and norms: these include spot-checks, supervisions (also relevant for M&E, see above) and **internal compliance inspections** to be undertaken regularly.

The social protection institution's watchdog can take the form of an **Inspectorate** – as is the case in South Africa, for example. Its aim is to: investigate the integrity of social assistance's legal and operational frameworks and ensure that regulatory frameworks are upheld; conduct internal financial audits and audits of the implementing body to ensure that laws and policies are adhered to; and investigate fraud, corruption and mismanagement within the implementing body.

5.2 EXTERNAL MEASURES

5.2.1 Tribunals, Courts and Ombudspersons

Some countries (such as South Africa) have developed **independent tribunals for social assistance** that can re-examine the decisions of the implementing body. These tribunals also disseminate standardized rules and regulations of social assistance and interpret legislation.[25]

A similar role can also be played by existing **courts** - an expensive and slow, yet important option for independent redress[26]. ILO Convention 128 of 1967 inscribes the right of any applicant to adequate professional representation or support before the courts.[27]

Ombudspersons typically exist outside the formal bureaucratic apparatus and have an advisory role; their recommendations are non-binding but are generally respected. Ombudspersons are most effective when referral procedures and up to higher escalation are well publicized and understood.

25 paralegaladvice.org - http://www.paralegaladvice.org.za/docs/chap07/01.html

26 The following considerations should be kept in mind when adopting courts as a redress option: a) Legal decisions may create negative externalities if the overall impact of decisions on the system is not considered; b) Use of courts as the primary grievance mechanism can create huge backlogs, politicisation and inefficiency; c) Use of courts can be regressive, as the poor and marginalised have less access to the court system. Public defenders, NGOs and CSOs can help to bridge this gap.

27 ILO, 1998, p. 124

5.2.2 Financial Audits

Financial audits can provide oversight of financing, with the goal of warranting the appropriate use of funds and improving management. Typically, an audit aims to ensure that: disbursement procedures and systems are correct; accounting records are maintained and updated regularly; internal control systems are adequate regarding payments, authorizations and reconciliation; and that expenditures comply with budget provisions.[28] Compliance and internal auditing units and inspection authorities should be independent. They monitor compliance with laws and fight social grant fraud, as well as investigate cases of financial or service mismanagement and criminal acts within the institutions.

The **national auditor-general** should conduct independent audits of the financial and account statements of the institution. Annual budgets, corporate plans, annual reports and audited financial statements should be prepared and submitted according to the applicable legislation. Annual Audit Coverage Plans could be used to inform the institution about auditing schedules, and to ensure that audits are regularly conducted. An internal audit unit could also be set-up.

5.2.3 Social Audits

Social audits have become increasingly popular as a tool to allow recipients and communities to review and provide feedback on programme implementation. This requires building and strengthening local capacity for the conduct of social audits and provides an opportunity to include vulnerable groups such as women and people with disability – or their representative organisations – in the process. Audits are often conducted by a local civil society organisation, which reviews process compliance, and operational rules and regulations, and presents results in a public hearing[29]. For example, India's MGNREGS public works programme has legally mandated the use of social audits, which have been successfully used to improve programme implementation (see Box 30).

Box 30: Social Audits in India's MGNREHS

> The first step in conducting the social audit is a notification to the relevant sub-district office regarding RTI (Right to Information) obligations and requesting unrestricted access to relevant MGNREGS documents. A team comprising of state and district auditors will, upon their arrival in the sub-district headquarter, first recruit and then intensively train village social auditors. In each Gram Panchayat, the social audit team verifies official labour expenses by visiting labourers listed in the worksite logs. Then a sub-district level public hearing is held with implementing officials to discuss the audit findings. Here, complaints will be read-out, testimonies verified and accused officials given an opportunity to defend themselves. After the public hearing, a Decision Taken Report (DTR) is created in which the responsibility for each confirmed wrong doing is pinned on a programme functionary (or on multiple functionaries).

Source: Ayliffe, Aslam and Schjødt (forthcoming)

5.2.4 Integrity Frameworks and Anti-fraud Campaigns

Fraud can be understood as the misrepresentation of interests to gain access to benefits illegally It is mostly ascribed to immoral and unethical behaviour. **Anti-corruption and fraud efforts should be integrated into operations and be among the strategic functions of the organization.** These may include:

- Developing an internal integrity framework, which should be based on clearly defined principles, such as honesty, fairness in operations, transparency, accountability, gender equality and inclusiveness, and disability inclusiveness; the framework needs to be available for everyone interested and should be a compulsory part of staff training and sensitization[30]
- Prevention measures include the production and dissemination of 'Principles and Values' statements stipulating expected behaviour, which should guide public servants in any professional activity or role
- Proactive forms of engagement including public anti-fraud campaigns launched with the support of high-profile actors/ champions, such as the presidency

28 Bassett, Giannozzi, Pop, and Ringold, 2012, p. 56
29 Subbarao et al., 2012.30 Public Service Commission, p. 26
30 Public Service Commission, p. 26

All of the above[31] should:

1. Give clear ethical standards for public sector employees
2. Give ethical guidance to public employees
3. Inform public servants about their rights and duties when detecting offences
4. Make clear that political commitment supports the ethical behaviour of public employees
5. Expect managers to demonstrate and promote ethical conduct
6. Promote ethical behaviour and conduct through management policies, procedures and practices
7. Promote ethical conduct through public service conditions and management of human resources
8. Provide adequate accountability mechanisms
9. Provide for decision-making processes that are transparent and open to scrutiny
10. Outline clear guidelines for collaboration between the public and private sectors, and with civil society organisations
11. Posit that misconduct should be sanctioned and solved through appropriate procedures.

5.2.5 Scrutiny by Parliament (e.g., Public Accounts Committee)

Sectoral Committees may be appointed within the government to scrutinize spending and delivery within the social protection sector. A cross-cutting example is the creation of a Public Accounts Committee responsible for overseeing government expenditures, and to ensure they are effective and honest.

5.2.6 NGOs and Media: Stimulating Demand for Accountability

The role of NGOs and media is important in facilitating and 'stimulating' access to complaint and appeal mechanisms, providing independent oversight and shoring up citizens' rights to entitlements and quality services.

NGOs and CSOs can sponsor and facilitate access to a redress procedure by aggregating or making claims for collective redress, providing information to citizens and providing legal aid to help access courts and legal redress mechanisms.[34] CSOs and NGOs can also play an important role in increasing the 'rights consciousness' of people who might not think they are entitled to a decent service or complain.[35] Similarly, NGOs and CSOs, including Organisations of Persons with Disabilities, may team up with the government to support participatory monitoring and evaluation of programmes, for example by becoming members of the Evaluation Reference Group, steering and advisory committees etc. and ensuring that these accountability mechanisms are asking the right questions, measuring the right indicators and reporting back to vulnerable groups, such as women and children and people with disability.

In Chile, a private foundation is instrumental in aggregating and publicising common complaints about public services for use by journalists and other NGOs. In the UK and the Dominican Republic, NGOs have been successfully 'contracted' and formed by the government to fulfil their role more effectively. In Sub-Saharan Africa, Helpage International has been active in several countries fostering Older-Citizens Monitoring - involving older people at the grassroots level to monitor the implementation of policies and services that affect their lives, and using the evidence to advocate for change.

The media provide a two-way route to enhance the accountability of social protection provisions. The media can be used by service providers to disseminate information on programme entitlements and existing Complaints and Appeal Mechanisms. An independent media can also serve an important watchdog function by publicising significant failures and complaints and appeals in service delivery and generating pressure to redress them.

However, the media can magnify news-worthy aspects of complaints and appeals, rather than what is primarily in the public interest. In Brazil, for example, publicity of fraudulent attainment of benefits is frequent and seen as 'scandalous.' Therefore, partly as a result of this, policymakers are far more likely to anticipate 'inclusion errors' in policy design rather than 'exclusion errors', which attract less attention.[36]

31 Pope, 2000, pp. 178

5.3 TAKE-AWAY LESSONS

- Most programmes have established some form of complaints and appeals mechanisms at community levels or more formally through tribunals, courts and ombudsmen. These play an important role in ensuring the accountability of the system. Other accountability mechanisms include spot checks, supervision, inspection and social audits.
- Accountability systems for social protection programmes serve the dual objective of a) ensuring and improving quality service delivery, and b) enhancing the impact of social protection provision.
- A wide variety of internal and external measures can be set up to increase the overall accountability of service provision in the social protection sector, in combination with a strong M&E system.

6

BIBLIOGRAPHY

Ackoff, R. L. (1989) "From Data to Wisdom", Journal of Applied Systems Analysis, Volume 16, p 3-9.

Andrews, M. Pritchett, L. and Woolcock M. (2017) "Building State Capability: Evidence, Analysis, Action, Oxford University Press". Available at: http://bsc.cid.harvard.edu/files/bsc/files/bsc_book.pdf

Attah, R., Barca, V., MacAuslan, I., Pellerano, L., Ragno, L.P., Riemenschneider, N. and Simon, C. (2015) "How to move beyond the impact evaluation trap? Challenges and solutions for the setting up of comprehensive M&E systems for Social Safety Nets." Oxford: Oxford Policy Management.

Bamberger M. (1991) "The politics of evaluation in developing countries", Evaluation and Program Planning.

Bastagli, F., Hagen-Zanker, J., Harman, L., Barca, V., Sturge, G., Tanja, S. and Pellerano, L. (2017) "Cash transfers: what does the evidence say? A rigorous review of programme impact and of the role of design and implementation features", Overseas Development Institute, London.

Bedi T. et al (2006) "Beyond the Numbers, Understanding the Institutions for Monitoring Poverty Reduction Strategies", World Bank publication.

Boaz, Ashby and Young, (2002) "Systematic Reviews: What have they got to offer evidence based policy and practice?" ESRC UK Centre for Evidence Based Policy and Practice: Working Paper 2.

Briceno, B. (2010) "Defining the type of M&E system: Clients, Intended Uses and Actual Utilisation", Nuts and Bolts of M&E systems special series, N. 2, September.

Chen H.T. (2005) "Practical programme evaluation: assessing and improving planning, interpretation and effectiveness, Sage Publications.

Clarke, A. (1999) Understanding Evaluation In: Evaluation Research, Sage Research Methods.

CLEAR (2013) "Demand and Supply: Monitoring, Evaluation, and Performance Management Information and Services In Anglophone Sub-Saharan Africa A Synthesis of Nine Studies", the CLEAR Initiative. Available at: https://www.theclearinitiative.org/ sites/clearinitiative/files/2016-04/Demand_and_Supply_Anglophone_Africa_2013.pdf

Davies, Nutley, Smith, (2000) "What Works? Evidence-Based Policy and Practice in Public Services", The Policy Press.

Davis B. (2003)"Innovative Policy Instruments and Evaluation in Rural and Agricultural Development in Latin America and the Caribbean in Current and Emerging Issues for Economic Analysis and Policy Research (CUREMIS), FAO.

Davis, Handa, Hyper, Winder Rossi, Winters and Yablonski, eds. (2016) "From Evidence To Action – The Story of Cash Transfers and Impact Evaluation in Sub-Saharan Africa", Oxford University Press.

Deaton A. and Cartwright N. (2016), "Understanding and Misunderstanding Randomized Controlled Trials", NBER Working Paper No. 22595.

Gorgens M., Kusek J.Z. (2009): "Making monitoring and evaluation systems work: a capacity development toolkit", World Bank publication.

Gosling l. and M. Edwards (1995) "Toolkits: a practical guide to assessment, monitoring, review and evaluation", Development Manual, Save the Children, UK.

Government of Ghana (2013) "Monitoring and Evaluation Framework and System Document (2013–2016), Livelihood Empowerment against Poverty Programme (LEAP)", Department of Social Welfare.

Government of Mozambique (2014), "Manual for Monitoring and Evaluation the National Institute of Social Action (INAS) programmes".

Grosh M. E. et al (2008) "For protection and promotion: the design and implementation of effective safety nets"; The World Bank, Washington, D.C.

Hatry, H. (2010) "Key steps in Designing and Implementing a Monitoring and Evaluation Process", in World Bank PREM notes, Nuts and Bolts of M&E systems special series, N. 5, December 2010.

Haynes L. et al (2012) "Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials", UK Cabinet Office, Behavioural Insights Team.

ILO (2012), Recommendation 202 on National Floors of Social Protection.

Kidd, S., Wapling, L., Bailey-Athias, D., & Tran, A. (2018). "Social Protection and Disability in South Africa. Development Pathways", Working Paper July 2018. Last accessed on 11 March 2022 at <https://www.developmentpathways.co.uk/wp-content/uploads/2018/07/Social-Protection-and-Disability-in-South-Africa.pdf>

Kidd, S., Wapling, L., Schjoedt, R., Gleders, B., Bailey-Athias, D., Tran, A. & Salomon, H. (2019). "Leaving No-one Behind: Building inclusive social protection systems for persons with disabilities", Development Pathways. Working Paper February 2019. Last accessed on 11 March 2022 at <https://www.developmentpathways.co.uk/wp-content/uploads/2019/04/Disability-overview-report-2019April07.pdf>

Kleinfeld, R. (2015), "Improving Development Aid Design and Evaluation: Plan for Sailboats, Not Trains", Carnegie Endowment for International Peace.

Kusek J.D, R. C. Rist (2004) "Ten steps to a results-based monitoring and evaluation system: a handbook for development practitioners", World Bank publication.

Kuster C. et al (2011) "Making evaluations matter: A practical guide for evaluators. Centre for Development Innovation, Wageningen University & Research centre, Wageningen, The Netherlands.

Mackay, K. (2007) "How to build M&E systems to support better government." Washington, DC: World Bank.

OECD (2002), "Glossary of Key Terms in Evaluation and Results Based Management", available here: http://www.oecd.org/ development/peer-reviews/2754804.pdf.

Palumbo D. (1987) "The Politics of Program Evaluation", American Journal of Evaluation, Vol 09 pp 54-56

Parkhurst (2017) "The Politics of Evidence: From evidence-based policy to the good governance of evidence", Sage Publication. Patel L. (2011) "Child Support Grants in Successful Social Protection Floor Experiences Volume 18: UNOSSC".

Pellens T. (2007), "Zambia National Social Protection Policy (NSPP) Monitoring and Evaluation (M&E) Framework - Evaluation Plan", report prepared by Oxford Policy Management for the Ministry of Community Development and Social Services, Government of Zambia.

Pritchett, Samji, Hammer (2013) "It's All About Me: Using Structured Experiential Learning ("e") to Crawl the Design Space", Harvard Kennedy School and Center for Global Development, Faculty Research Working Paper Series.

Rubio, G. (2011) "Results Readiness in Social Protection & Labour Operations: Technical Guidance Notes for Social Safety Nets Task Teams". SP discussion paper, World Bank.

Samson et al (2006) "Designing and Implementing Social Transfer Programmes" Economic Policy Research Institute. Capetown.

Seekings J. (2016) "Affordability and the political economy of social protection in contemporary Africa" CSSR Working Paper No. 389 Centre for Social Science Research.

Segone M. (2008) "Enhancing Evidence Based Policy Making through Country Led Monitoring and Evaluation Systems.", in Country- led monitoring and evaluation systems Better evidence, better policies, better development results, UNICEF.

Seidenfeld D and Sudhanshu Handa (2011) Zambia's Child Grant Program: Baseline Report 28 November, 2011 Final. American Institutes for Research Washington, D.C. USA.

Simon, C. Pellens, T. and Marzi, M. (2016) "Zambia National Social Protection Policy (NSPP) Monitoring and Evaluation (M&E) Framework", report prepared by Oxford Policy Management for the Minstry of Community Development and Social Services, Government of Zambia.

Thome K, Taylor E et al (2016) "The Local Economy Impacts of Social Cash Transfers: A Comparative Analysis of Seven Sub Haran Countries". Food Agriculture Organisation. Rome.

UNDP (2009) "Handbook on Planning, Monitoring and Evaluation for Development Results". UNICEF (2009) "Country-led monitoring and evaluation systems".

Vaessen J. (2010) "Challenges in Impact evaluation of Development Interventions: Opportunities and Limitations for Randomized Experiments." Institute of Development and Policy and Management.

Weiss C. (1993) 'Where Politics and Evaluation Research Meet" Evaluation Practice, Vol 14, No 1 , Available at: http://www.sfu. ca/~palys/Weiss-1973-WherePoliticsAndEvaluationMeet.pdf.

Weiss Carol (1970) "The Politicisation of Evaluation Research" in Journal of Social Issues 26 (4) 57-86

White H. and Sabarwal S. (2014) "Quasi Experimental Designs" Methodological Briefs Impact Evaluation No 08 UNICEF.

World Bank (2010) "Safety Nets How to, Monitoring and Evaluation", available at: http://siteresources.worldbank.org/SAFETYNETSANDTRANSFERS/Resources/281945-1291746977764/10-me.pdf.

WWP (2016) "Bolsa Familia Interfederative Coordination Instruments: the IGD, World Without Poverty Initative", available at http:// wwp.org.br/wp-content/uploads/2016/12/igd_presentation-interfederative-coordination-instruments.pdf.

CURRICULUM **OVERVIEW**

The TRANSFORM Learning Package
is organized in a modular structure, and reflects the key building blocks of a holistic & interdependent social protection system.
The TRANSFORM modules that are currently available are listed below.
Other modules are under development and will be added to the curriculum.

LEG	**Legal Frameworks**
S&I	**Selection & Identification**
ADM	**Administration and Delivery Systems**
COO	**Coordination**
GOV	**Governance, Institutions & Organizational Structure**
MIS	**Management Information Systems & Approaches to Data Integration**
FIN	**Financing & Financial Management**
M&E	**Monitoring & Evaluation**

All TRANSFORM materials are available at:

www.transformsp.org

WHAT IS TRANSFORM?

TRANSFORM is an innovative learning package on the administration of national social protection floors in Africa. The prime objective of TRANSFORM is to build critical thinking and capacities of policy makers and practitioners at national and decentralized levels to improve the design, effectiveness and efficiency of social protection systems. TRANSFORM aims not only at imparting state-of-the-art knowledge that is appropriate for the challenges faced by countries in the region, but also to encourage learners to take leadership on the change and transformation of nationally defined social protection systems.

WHY TRANSFORM?

Many training curricula exist in the field of social protection and thus fundamental ideas, concepts, approaches and techniques are accessible. And yet, institutions and individuals struggle with the complexity of developing a broad, encompassing social protection system.

This complexity requires a transformational approach to teaching and knowledge sharing. It is far from enough to impart knowledge, to fill heads. It requires learners to grapple with the features of complexity, to stimulate creativity, to appreciate diversity and uniqueness, to be involved as a key element of ownership –elements which are at least as important as the factual knowledge itself. This learning package aims at just that: TRANSFORM!

Contact the TRANSFORM initiative at: **transform_socialprotection@ilo.org**
or visit **www.transformsp.org**

IMPLEMENTED, SUPPORTED AND DRIVEN BY:

African Union

ECOWAS CEDEAO

OECD DEVELOPMENT CENTRE dev — Finland — EU Social Protection Systems Programme

Food and Agriculture Organization of the United Nations

FRIEDRICH EBERT STIFTUNG

International Labour Organization

ITC International Training Centre

IngeniousPeoplesKnowledge — Why not unleash ingenuity within systems?

Irish Aid — Rialtas na hÉireann — Government of Ireland

LEARN TO LEAD — imagine learning differently

SASPEN — Southern African Social Protection Experts Network

socialprotection.org

UKaid — from the British people

UNDP

unicef

UN WOMEN

UNITED NATIONS UNIVERSITY — UNU-MERIT

WFP World Food Programme