

Beyond building blocks?

Identifying and monitoring dynamic drivers of sector performance.

Synthesis report, March 2019.



Cover images: (main) Nawoli Jesca, Commercial Officer, and Nkundizana Julius, Team Leader, Busolwe Piped Water Supply System, checking on a pipe to the main water reservoir, Busolwe Piped Water Supply System, Butaleja District, Uganda; (top) The new toilet block at Kebessa Elementary School, Burie district, Amhara, Ethiopia.

Acknowledgements

This report draws on the findings of WaterAid-funded research conducted by the Overseas Development Institute (ODI). It was written by Nathaniel Mason, Fiona Samuels and Miriam Denis Le Seve, from ODI, with input from colleagues Guy Jobbins and Leni Wild. The authors thank the many experts interviewed for the study. The views presented in this publication are those of the authors and do not necessarily represent the views of ODI or WaterAid.

WaterAid would also like to thank Stuart Kempster, Helen Hamilton, Henry Northover, Ian Gavin, Ellen Greggio and Alison MacIntyre for their contributions in reviewing this work.

For more information please contact:

Clare Battle

Senior Policy Analyst, WaterAid

clarebattle@wateraid.org.uk

+44 (0) 20 7793 4502

Beyond building blocks?

Identifying and monitoring dynamic drivers of sector performance: Synthesis report

Executive summary	3
1 Introduction	5
2 Building blocks in WASH and health: Experience and risks	8
2.1 Evolution of building block frameworks in the WASH sector	8
2.2 Learning from building blocks in the health sector	9
2.3 Four risks arising from monitoring with building block frameworks	10
3 Addressing the risks of monitoring systems with building blocks and other approaches	12
3.1 Incentive risk 1: Isomorphic mimicry	12
3.2 Incentive risk 2: Siloes, not system	15
3.3 Incentive risk 3: Overlooking emergence	16
3.4 Incentive risk 4: Misplaced effort	18
4 Monitoring sector systems in practice: Country case studies	20
4.1 Indonesia's WASH sector	20
4.2 Ethiopia's WASH sector	23
4.3 Ethiopia's health sector	26
4.4 Addressing the incentive risks in country-level system monitoring	29
5 Ways forward: Conclusions and recommendations	30
5.1 Conclusions	30
5.2 Recommendations	35
Annex 1: List of organisations consulted	37
Annex 2: Timeline of international building block initiatives in WASH	38
Annex 3: Principle themes in international WASH building block frameworks	41
References	42

Executive summary

A system-wide approach is essential to achieve the Sustainable Development Goal (SDG) targets of universal and equitable access to safe water, sanitation and hygiene (WASH) for all. The right metrics, applied in the right way, can help track progress towards these targets and incentivise system strengthening behaviour among donors, governments and implementing partners. Enhancing current approaches to monitoring systems, to capture important drivers of system performance in a dynamic, complex sector, is therefore critical to accelerate progress towards the SDGs.

To date, most approaches proposed for monitoring WASH and health sector systems are based on ‘building blocks’ – elements or sub-systems that are usually defined around broad functions, such as policy setting, financing, capacity development, and accountability. These approaches have provided a structure for discussing the different actors, roles, resources and institutions at work in a sector system, and the related challenges or bottlenecks.

However, practical applications of building block frameworks for ongoing, routine monitoring have been limited. So far, they have more often been used to support one-off diagnostic or review exercises. And experience from the WASH and health sectors suggests that when these frameworks are used for monitoring sector systems there are four key risks:

1. Building blocks, and indicators used to measure them, encourage a focus on what a sector system should look like. But a **focus on ‘form’** may not say much about, or incentivise, improved ‘function’ within the system.
2. A building block approach involves dividing up the WASH system into more manageable component blocks. This may **overlook the interactions** between different sub-systems and across governance levels that are vital drivers of sector performance, such as learning, coordination and political commitment.
3. Building block frameworks can encourage a **static view** that does not adequately recognise, capture or measure unpredictable changes, which are often a feature of complex systems.
4. Most building block frameworks aspire to be comprehensive, but this does not necessarily help tackle complex problems. A **lack of focus** on the binding constraints on system performance leads to undue attention on marginal issues and encourages over-complex responses.

Several building block frameworks, as well as adapted or alternative approaches to monitoring systems change, acknowledge these risks and point to possible responses. These range from efforts to frame indicators that stress institutional function rather than form and develop supplementary tools that focus on systems interactions and relationships, to more drastic responses such as problem-driven iterative adaption.

At country level, Indonesia has used results-based funding mechanisms to refine indicators that can incentivise system strengthening, as well as triggering payment, but this has also shown the challenges in doing so. Ethiopia's WASH and health sectors provide examples of emerging indicators aimed at monitoring system strength, but also highlight the possibility for system strengthening efforts to be sustained in the absence of robust routine monitoring of system performance.

Analysis of these different approaches leads to the following recommendations:

1) WASH donors and implementing partners should continue to use building block frameworks, but with greater emphasis on using and improving them for recurrent monitoring of sector performance.

2) To address the possible risks and limitations of monitoring systems with building blocks, donors and implementing partners should work together on the following priorities:

- **Carry out research to build understanding of the core building blocks that capture, at a high level, the key ingredients or binding constraints for performance of a given type of WASH sector or subsector system.**
- **Empower local stakeholders to define the detailed indicators for monitoring progress in each of these core building blocks, as well as any additional, locally-specific building blocks, according to the context.**
- **Fund small-scale reform 'experiments', backed by monitoring with rapid feedback, to iteratively sharpen the focus on what should be prioritised for reform and further monitoring.**
- **Select indicators to track how institutions function, rather than just what they look like – ranging from quantitative proxies for performance such as non-revenue water, to measures of user satisfaction.**
- **Further develop, test and simplify methods to capture and track system-level dynamics and interactions, for example, approaches that map and measure the strength of networks between actors or relationships between issues.**
- **Leave space to capture unpredicted, emergent issues within results frameworks, balanced with ground rules to define what counts as significant.**
- **Promote recurrent reviews of each system monitoring framework, to ensure they are kept as relevant as possible to the system in question.**

These predominantly technical responses will be a move in the right direction, but greater political awareness and agility is also needed. Governments and development partners must exploit whatever political space they have, to look beyond narrow definitions of results that reduce everything to infrastructure or anonymous beneficiaries. They must tell human-centred stories about how stronger WASH systems make a real difference to people. And they must continue to foster a culture of learning and evidence-informed decisions based on monitoring data.

1 Introduction

The Sustainable Development Goal (SDG) targets on water supply, sanitation and hygiene (WASH)¹ will not be achieved without a system-wide approach.

Governments must be supported to strengthen the core constraints in the sector, so they can deliver and sustain services to all citizens. Compared to the Millennium Development Goal (MDG) targets, the SDG WASH targets require a much higher level of service, for everyone, that is sustained over time. These dimensions of the SDG WASH targets – quality, universality and sustainability – demand a step-change in how donors and governments approach WASH.

At the same time, a move to this system-wide approach is both practically and politically challenging; nowhere more so than in how progress is measured and communicated. Throughout the MDG period, an emphasis on first-time access to improved water supply and sanitation facilities provided a simple measure of progress. Against this, donors and governments had a practical way to measure their contribution, and it was politically straightforward to communicate that contribution to the tax-paying or donating public.

However, the MDG WASH access targets and indicators also created some harmful incentives, including biasing funding towards new infrastructure and the easiest people to reach. SDG targets 6.1 and 6.2, and their corresponding indicators, attempt to adjust these incentives, by requiring universal access to safely managed WASH services. However, it can take years for an increase in finance, or a policy reform, to translate into WASH services. To tell whether a WASH sector is going in the right direction, it is necessary to track changes in the service delivery system itself, not just the services it eventually delivers. And, all the while, donors and the governments they fund face increasing demand for robust and easily communicable results of their aid programmes.

To date, most efforts to understand weaknesses across an entire WASH sector system – whether at national or more local levels – have taken ‘building blocks’ as their organising principle. For the purposes of this paper, we adopt the definition proposed by IRC, which focuses on the production of key functions for services, and acknowledges the role of both actors and factors (human and non-human elements) in a WASH system:

A building block is a recognisable (and widely recognised) sub-system within the larger WASH system, whose actors and factors work together to perform a key function.²

Definitions used in this report

Throughout this report, the term ‘WASH system’ is used to refer to the entire WASH sector, which is understood to involve multiple levels, stakeholders, relationships and interactions, both within the sector and in relation to the broader context. Some WASH stakeholders may think of a physical water or sanitation network when hearing the term ‘WASH system’ – as did several WASH representatives whom we interviewed. However, we feel the wider utility of ‘systems thinking’, and concepts such as ‘system strengthening’ in fields including health, justify using the term in this more encompassing way.

By ‘monitoring WASH systems’ we mean a deliberate, regular process in which defined aspects of the system are tracked, enabling an assessment of change over time and/or comparison with other systems that might be monitored in a similar way.

Monitoring WASH systems is distinct from monitoring WASH outputs and outcomes – the infrastructure or services that the system is expected to deliver. A distinction can also be made between monitoring a WASH system and diagnosis or assessment of the system, which might be undertaken as a one-off exercise to inform planning, as well as periodic review of the system, in which a changing set of issues are discussed. Examples of the former could include a bottleneck analysis, while joint sector reviews often fall into the latter category.

As this report makes clear, building block frameworks have played, and continue to play, an important role in supporting WASH system strengthening efforts: particularly to inform policy and investment decisions through one-off analyses. To a much lesser extent, they have also informed ongoing monitoring of progress, whether at the level of international policy (e.g. UN-GLAAS), national and international commitments (e.g. SWA), or individual aid programmes and projects.

However, this report argues that dividing a complex sector system into constituent building blocks overlooks some of the most important aspects of its behaviour as a system. In turn, this can create incentives that do little to help donors and governments to strengthen the system. We identify four key risks of using building block frameworks to assess and monitor WASH systems: (i) focusing attention on the presence of certain, prescriptive institutional forms, rather than how they actually function; (ii) emphasising individual elements, rather than interactions within the system; (iii) being insensitive to the emergence of unforeseen issues; and (iv) encouraging comprehensiveness, rather than selectively identifying which factors matter most for the performance of the system.

The report also finds that while these are justifiable concerns, it is not clear that any single alternative approach would address all four risks effectively. Instead, it is likely that elements of different approaches can be combined to make building block approaches more effective, to monitor and incentivise system strengthening.

To arrive at this conclusion, the report takes stock of existing and emerging frameworks and approaches to assess WASH systems and measure changes in them. This includes a range of building block type frameworks, but also alternative and adapted approaches used by donors and recipient governments to track and incentivise system-level change – including results-based funding and periodic sector review mechanisms. Across these different approaches, we assess how drivers of change are identified, progress is monitored, and further progress incentivised. We focus on how donors monitor systems, and the incentives this creates for them and the governments they provide funding to. However, using country case studies, we also consider how donor and government monitoring of sector systems might evolve together, creating a more complex web of incentives for both parties.

The report aims to support WASH sector stakeholders focused on sector performance – first and foremost, development partners and global and regional initiatives, but also developing country governments – to maintain financial and political attention on the key drivers of sector performance. Four key questions guide the analysis throughout, and we return to them in the conclusion:

1. How is sector/system strengthening currently being monitored by donors, governments and the international community?
2. What sort of behaviours, activities and programmes does this incentivise?
3. To what extent does this align with what we know about complex and dynamic nature of systems change?
4. What innovations in monitoring complex systems could be usefully applied to maintain a focus on the key drivers of sector performance in a dynamic WASH sector?

The report builds on a previous study³ and is based on qualitative research, including a desk study and key informant interviews (see Annex 1) carried out between October 2018 and January 2019. Based on the sum of the evidence collected, including two case studies (Indonesia and Ethiopia) and sectors (WASH and health), it offers a set of core principles and recommendations.

The rest of the document is structured as follows. In Section 2, we outline the evolution of frameworks for monitoring system strengthening in WASH and health, which have principally used building blocks as their organising principle. Through this analysis, we explain in more detail the four key risks of creating perverse incentives, which arise from using a building block lens to monitor the strengthening of complex and dynamic systems. Section 3 considers how building blocks, as well as adapted or alternative approaches, have tried to manage these risks, focusing on approaches originated by donors and international organisations. To ground this analysis in the experience of specific sectors, in specific countries, Section 4 looks at how system monitoring has been approached in the WASH sector in Indonesia and Ethiopia, as well as in Ethiopia's health sector, and the incentives at work. Here, the focus is on approaches that have co-evolved between government and donors/international

organisations. In Section 5, the report concludes with an overarching assessment of the evidence, and a set of recommendations.

2 Building blocks in WASH and health: Experience and risks

2.1 Evolution of building block frameworks in the WASH sector

Building block frameworks applied to WASH have evolved for over a decade (see Annex 2 for a timeline). Three major international initiatives have done much to popularise building block frameworks as a way to understand and, to a lesser extent, monitor WASH systems: the World Bank's Country Status Overview (CSO) series; the UN-Water Global Analysis and Assessment of Sanitation and Drinking-water (GLAAS), implemented by the World Health Organization (WHO); and UNICEF's WASH Bottleneck Analysis Tool (WASH BAT).

Each of these has a slightly different selection of building blocks or constituent sub-criteria, and they also differ in terms of how building blocks are grouped and themes therefore prioritised (see Annex 3 for an assessment of common and distinct themes).ⁱ This reflects the fact that each framework is, to a large extent, a normative product, reflecting the perspectives of a limited group of people. There are numerous overlaps, however, given extensive dialogue and joint-working between the leading agencies. More recently, a set of five building blocks were adopted by the Sanitation and Water for All (SWA) partnership, comprising sector policy & strategy; institutional arrangements; sector financing; planning, monitoring and review; and capacity development.⁴ In addition to the frameworks mentioned above, the SWA building blocks also build on research and proposed frameworks developed by other partners, including WaterAid⁵ and IRC.⁶

The international community may be moving towards consensus on the core WASH building blocks. For example, the SWA framework now underpins UNICEF's second version of the WASH BAT tool. Nonetheless, other agencies continue to develop their own building block frameworks, including Water for People,⁷ IRC,⁸ WaterAid⁹ and Water and Sanitation for the Urban Poor (WSUP).^{10,ii} This continued evolution reflects the intuitive appeal of a building block framing to analyse sector systems, and the increasing interest in systems-oriented approaches in WASH, more generally. Donors also continue to signal their interest in this space – several have released WASH-related strategy documents with system strengthening at the centre.¹¹ Our interviews confirmed that several donors are developing monitoring approaches to accompany their increased focus on systems change.

ⁱ The GLAAS initiative tends not to use the term 'building block', though the questions included in its biennial survey to country governments on different aspects of the WASH system overlap with the questions used in the CSOs and WASH BAT, and they are organised according to similar thematic categories.

ⁱⁱ We highlight some of the more recent examples. For an older but more comprehensive assessment, see Schweitzer et al (2014), which looks at frameworks and tools through the lens of sustainability rather than systems in general, but covers many building block type frameworks used at sector, project and other scales.

However, the newer WASH building block frameworks released in 2017 and 2018 reflect the difficulty in establishing agreement about what the essential building blocks are, how they combine to constitute a WASH system, and how they should be measured. While the newer frameworks may prioritise a few new issues (both WaterAid and IRC's most recent frameworks,^{12,13} for example, have water resources management as a dedicated building block) they generally acknowledge and build on existing models and tools. Their intended application is somewhat different too; most are intended to be used at a more local level (for example, a district or municipality) as well as, or instead of, the entire national WASH system.

Across most of the frameworks, a crucial observation is that they have not yet been applied for recurrent *monitoring* of WASH system strength over time, at either a national or sub-national level. More often, they have been used as a diagnostic tool that provides an opening for policy dialogue or a basis for planning. While two rounds of CSOs were conducted in Africa, the framework was developed substantially, and more countries included for the second round, limiting the viability of monitoring progress over time. In some cases, such as IRC's framework, baselines have been established with the intention of conducting recurrent monitoring. UNICEF's WASH BAT has been used more than once in four countries (Madagascar, Nigeria, Pakistan and Vietnam) but the tool was developed between the first and second assessments. Pakistan initiated application of the new version of the WASH BAT annually at both national and subnational levels from 2017, and there is an intention to make recurrent assessments using WASH BAT in more of the 35 countries that have used the tool to date (with 10 more expected to do so for the first time in early 2019 – interview data). Other examples of monitoring using the same building block-type framework more than once, at a regular interval, include GLAAS at the international level, and Water for People's framework at the national and subnational levels – so far, the latter has been used in 2017 and 2018, in 26 districts across nine countries.

2.2 Learning from building blocks in the health sector

The importance of a systems approach is arguably longer established in the health sector, driven by, among other things, the need to move away from disease-specific approaches and to ensure that vulnerable groups benefit. The health sector has been using building block approaches to assess and monitor health system strengthening for some time, seemingly with more discussion of the risks and possible responses. Given the enthusiasm with which building block frameworks have been adopted in WASH, it is worth considering what lessons can be learned from health.

The most well-known and widely applied building block framework in health is that proposed in 2010 by WHO and its global partners. This defined a health system as all the organisations, institutions, resources and people whose primary purpose is to improve health.¹⁴ The framework comprises six core building blocks: service delivery; health workforce; health information systems; access to essential medicines; financing; and leadership and governance. Leadership and governance,

and health information systems, are framed as cross-cutting components that provide the basis for the overall policy and regulation of the other system blocks. Financing and the workforce are key input components, while medical products/technologies and service delivery reflect immediate system outputs. For each of the building blocks a set of core indicators was selected to measure progress, alongside a methodology to monitor the indicators.

Since its development, several challenges have been identified. A first is around the ability of building block approaches to capture dynamism and complexity. While the selection of indicators for the building blocks was guided by the need to detect changes and progress,¹⁵ it has been argued that there is insufficient attention to the properties of health systems as complex adaptive systems, made up of dynamically changing and interconnected elements.¹⁶ There is also insufficient attention paid to the fact that different elements in the systems may interact in different ways, according to different contextual factors or historical moments.¹⁷

Another criticism is that the building block approaches under-value important interrelationships between and within systems. By drawing a relatively narrow boundary around the health sector, they may miss interactions and linkages with other sectors (including WASH). Underlying social and economic determinants may also be missed, such as the influence of gender inequities and education on health behaviours and outcomes.¹⁸ Relatedly, they may not lend themselves to an understanding of system strength as being made up of interactions at different levels, for example, with respect to leadership and governance.

Finally, the *quality* of services that health systems can deliver has also emerged as an increasing concern, with the drive towards universal health coverage. While the WHO building block framework includes indicators for quality, a recent Lancet Commission proposes a much greater focus on the views of the people the system serves. This would require monitoring both user experience and user confidence in the system.¹⁹

There have been various attempts to address the shortcomings of building block approaches in the health sector, of which the proposals of the Lancet Commission are one example. Another is the work of the Health Systems Governance Collaborative to find more appropriate ways to conceptualise and measure governance. Proposals include assessing a range of measures for governance structures, processes and outcomes, at a range of scales from an individual organisation or facility to the wider political economy.²⁰ Broadly, however, the health sector has not yet progressed to mainstreaming these solutions into widespread practice.

2.3 Four risks arising from monitoring with building block frameworks

The lessons learned from the health sector suggest that, for all the enthusiasm for building block frameworks in WASH, there are certain risks with using them to understand and monitor sector systems. If building blocks miss certain features of

the system, or focus attention on unimportant issues, they could in turn incentivise investments, policies and programming that do little for, or even detract from, system strength. Based on this learning, as well as assessment of the many approaches in WASH (Annex 2), we identify four key risks whereby monitoring using building blocks could result in perverse incentives.

Risk 1: Encouraging isomorphic mimicry, by focusing on form over function

It is usually easier to assess whether something exists than how it performs. The risk, however, is ‘isomorphic mimicry’ – which ‘conflates form and function: “looks like” substitutes for “does”... Appealing budget documents count even if they don’t determine spending outcomes.’²¹ In the WASH sector, other examples could include a model sanitation policy document that is published, but not implemented, or a regulatory agency for water supply that is set up but unable to enforce regulation. Prescriptive building block frameworks that set out what the WASH sector should look like, according to international ‘best practice’, are vulnerable to conflating form and function. Being prescriptive about the ideal form for building blocks, through standardised indicators, can be desirable for donors who seek to aggregate and compare results across countries, or governments who want to do so between subnational jurisdictions. But as we shall see, it should be possible to design indicators so they better assess how the system actually functions and performs. It should also be possible to provide for comparability at a high level, while allowing the detail of specific metrics to be defined locally, recognising the unique web of behaviour and incentives that make up a WASH sector, as a social-political system.

Risk 2: Promoting siloed responses, by emphasising individual blocks rather than system-level interaction

By definition, a building block approach involves dividing up the WASH system’s complexity into more manageable component blocks. This is pragmatic, but intuitively at odds with the idea of complex systems being ‘characterised by interconnected and interdependent elements and dimensions’.²² Issues such as ‘learning’, ‘coordination’ or ‘leadership’ are a particular concern. These could be assessed by looking at individual entities or sub-systems which have a designated role for learning and leadership. However, both phenomena arise as the product of multiple interactions across all parts of the wider system. Research on political drive in the sanitation sector, for example, underscores how leadership needs to manifest across sectors and at different tiers of government, not just as high-level political commitment in a ministry or president’s office.²³

Risk 3: Overlooking emergent issues, by encouraging a static rather than dynamic view of the system

Change in complex systems is unpredictable, and action in one part of the system can have unexpected effects elsewhere. Pre-determined building blocks and rigidly fixed indicator sets can encourage stakeholders to view change as a linear process, rather than one which is ‘emergent’ (whereby ‘the behaviour of systems emerges –

often unpredictably – from the interaction of the parts'²⁴). As a key characteristic of complex systems, emergence is a subset of the preceding challenge, of monitoring the behaviour of a system as a whole, as opposed to its component parts. However, we give it special attention because of this study's interest in the ability of system strengthening monitoring to capture dynamism (see Research question 3, Section 1). As for Risk 1, there may be a balance to be struck between leaving room to capture unexpected but important phenomena and maintaining a base level of consistency to allow comparison, both between jurisdictions and over time.

Risk 4: Prioritising issues that are less important for, or even detrimental to, system strength, by privileging comprehensiveness at the expense of focus.

There can be a temptation to address complexity, or at least complication, by being comprehensive. Most attempts to describe a WASH system in terms of constituent building blocks have struggled to keep the total number of distinct, underlying indicators below 20. Variants and additions are also often introduced in recognition of the essential differences between WASH subsectors (urban/rural; water supply/sanitation/hygiene). Comprehensiveness is perhaps an advantage when building blocks are used to frame a one-off, discursive exercise to inform planning, but the time-cost adds up when regularly assessing indicators as part of recurrent monitoring and evaluation. More importantly, monitoring that does not identify which failures within a system matter most could paralyse reform efforts, lead to undue attention to marginal issues, or encourage over-complex responses. Tools to help understand and monitor complex systems should help to simplify, by identifying the core, binding constraints that really matter for system strengthening.

The above four risks are especially relevant given the general direction of aid spending across many sectors, including WASH. Several of our interviewees reflected on a prevailing trend back towards projectised approaches, focused on narrow, easily counted and attributable results, and away from more programmatic approaches to sector support. Without robust metrics for performance of WASH sector systems, there is a risk that the associated investments could be rolled back. Nonetheless, it is important to recognise that many building block frameworks already recognise and try to address these risks, as do adapted and alternative approaches that are emerging. We turn to these efforts in the next section.

3 Addressing the risks of monitoring systems with building blocks and other approaches

Using illustrative examples, this section considers how, and how far, the risks and corresponding incentive issues have been managed, by those advancing building blocks, as well as other adapted or alternative approaches.

3.1 Incentive risk 1: Isomorphic mimicry

The first and simplest response to this risk is to take more care in how building block indicators are identified and interpreted, to ensure attention to function rather than

form. This is the approach adopted by several existing building block frameworks. The second round of CSOs, for example, tried to capture policy implementation, if not the performance of that implementation, in assessment of each building block. This was done through qualitative indicators (for example, whether institutional roles are not only defined, but also operationalised) and more quantitative indicators (such as the percentage of budgets used).²⁵ IRC meanwhile recommend framing indicators that stress institutional function (e.g. regulation) rather than form (e.g. an independent regulatory agency – interview data).

It is still likely, however, that normative views about how the system *should* look, based on internationally accepted blueprints, can have an influence on what gets monitored and prioritised.

A second response would be to use complementary incentive mechanisms to ensure system monitoring is locally owned and grounded. A possible example is the Directorate-General for International Cooperation (DGIS) of the Netherlands' use of three complementary mechanisms to monitor and incentivise the sustainability of its WASH investments: the **sustainability clause, check and compact**. Although the emphasis is on sustainability, the three-part approach has a system strengthening ethos, insofar as it tries to encourage system-level interventions in support of sustainable services:

- The sustainability clause requires the implementer 'to ensure service delivery for 10 years after project completion'.
- The sustainability check is a periodic, mandatory assessment of service functionality, as well as the ability of the system to sustain services, carried out by an independent third party and requiring a management response.
- The sustainability compact is an agreement developed locally between the implementing agency and others, setting out 'the roles and responsibilities of local stakeholders, including local government, beyond the project period'.²⁶

The sustainability check includes what is essentially a building block assessment, in which monitoring of the system's ability to sustain services is assessed under financial, institutional, environmental, technological and social headings. Like other building block frameworks, there is little inherent to the design of the checks that prevents a focus on form over function, incentivising isomorphic mimicry.

By adding the sustainability clause and compact, however, DGIS appears to be making a more direct attempt to make the check, and subsequent follow-up, more relevant to the specific system and its functional strengths and weaknesses. Because the financial penalty in the clause can occur after considerable time, it may encourage more serious consideration of which locally specific system failures are most important, and what can be done to address them in collaboration with local stakeholders. If the penalty were set in the nearer term, the implementing party might be able to avoid it by plugging a gap in the system itself. Yet, this is less likely to provide a viable solution over the course of 10 years. The compact, meanwhile,

complements this effect, by encouraging local institutions to take ownership of the reforms needed to address system weaknesses, which could undermine sustainability.

These effects are somewhat theoretical, however, and there are various challenges to rebalancing the incentives in this way. The first issue is that while the clause has some financial and legal ‘teeth’, there are no cases to date where a financial penalty has been exacted on a grant recipient, for failure to ensure sustainability.ⁱⁱⁱ A second issue is that the compact lacks such teeth altogether, and ownership on the part of other signatories is therefore far from guaranteed. What’s more, even if there is some moral leverage, through the signing of the compact, this does not mean signatories take ownership of specific actions identified through the monitoring (checks). Nor will it have much effect on important non-signatories, such as subnational government. Both issues were identified in a review of UNICEF’s experience applying bottleneck analysis, compacts and checks (and implicitly, the clause) in eight West African countries under DGIS-funded WASH programmes.²⁷

A third, more drastic, response is to overhaul not only how system monitoring is undertaken, but how interventions seeking to reform systems are approached altogether. **Problem-driven, iterative adaptation** (PDIA), for example, calls for strategies ‘that begin with generating locally nominated and prioritised problems, and that work iteratively to identify customised “best fit” responses’.²⁸ Similar principles have been advocated under the banner of ‘adaptive development’ and ‘doing development differently’.²⁹ Focusing on the monitoring aspects, these approaches imply (i) a much tighter loop between implementation, monitoring and adaptation; (ii) a need to attempt, and monitor, several small-scale reform strategies in parallel, iterating continuously to discover ‘what works’; and (iii) a greater willingness on the part of donors to fund and measure over longer timescales. It is relatively early days for the application of such approaches, though there are significant investments in research and learning, such as the UK and US-funded Global Learning for Adaptive Management initiative.³⁰ We did not identify any examples of WASH programmes doing so explicitly, though some agencies have expressed their interest.

A final response to shape monitoring systems around locally nominated priorities and functional performance is to **invest in strengthening countries’ own monitoring and regulatory systems**, so that they themselves can identify and monitor metrics of system performance. Many donors have supported the development of country WASH sector monitoring and management information infrastructure and capacity. However, examples of this support leading to better *systems* monitoring, and of generating data that donors can themselves use to track and justify their investments, appear to be rarer.

ⁱⁱⁱ DGIS has reportedly so far found the ‘threat’ that implementing partners could lose their preferred partner status sufficient, rather than needing to enforce a legal/financial penalty. Remedial action has been encouraged in this way, when checks have revealed problems – for example, when sanitation facilities in Mozambique did not qualify as sustainable because fly covers were missing, this was addressed by the partner (interview data).

As an example, over two decades and in several countries, the German Corporation for Development Cooperation (GIZ) has supported national regulators or other agencies to carry out utility performance monitoring and oversight.^{31,iv} Some of the data generated by the supported country-level monitoring and regulatory systems can, in principle, be aggregated, to provide proxy measures of system strength – for example, levels of cost recovery among utilities. However, GIZ reportedly still retains parallel monitoring frameworks as a tool for project managers to steer investments, rather than relying exclusively on the data generated by the local regulator or monitoring system (interview data). The example also underscores that donors face different political constraints on their system strengthening activities, and what they need to monitor and report on to justify them. For example, Germany appears to have relative political freedom to make these kinds of investments over the long term, and may face less demand for regular, aggregated results – whether at output, outcome or system level – than some other donors. It is also still possible that the provision of technical assistance does itself involve promotion of best-practice blueprints about what the sector system, and indeed monitoring systems, should look like.

3.2 Incentive risk 2: Siloes, not system

Again, the simplest response to the risk of overlooking interactions, relationships and other system-level phenomena is to apply care when using building blocks for monitoring: ‘maintaining a whole systems perspective’, while using building blocks to simplify ‘to a practical level that can support action’.³² Software could help this endeavour to some extent; for example, UNICEF’s WASH BAT online tool allows the identified causes of a given bottleneck to be replicated and assigned to other bottlenecks.³³ One step further could be to assign specific building blocks to capture some of the more ‘system-level’ phenomena. IRC, for example, includes a building block on ‘learning and adaptation’ in its framework, which assesses processes such as joint sector reviews and interaction between stakeholders at different levels (interview data).

We identified limited convincing examples that go beyond this. Proponents of applying complexity theory to international development point to the importance of complex system characteristics that can only really be understood at the system level, such as interdependence, feedback loops, non-linear behaviour, self-organisation and co-evolution.³⁴ Yet, the development community, not only in WASH, is still grappling with how to operationalise these concepts and find ways of measuring ‘systemic’ characteristics and behaviour over time.

The most promising examples originate with USAID, which has been developing conceptual and operational tools for understanding and monitoring complex systems

^{iv} GIZ is not a donor, but implements projects funded by the Federal Ministry for Economic Cooperation and Development (BMZ), often in partnership with the grant and soft-loan provision of the German development bank, KfW.

for some time.³⁵ A four-year strategic programme on analysing complexity and evaluating complex systems (SPACES-MERL) comes to an end in 2019.^v Its 2016 white paper on systems and complexity identified 24 systems tools and approaches for monitoring, evaluation and learning, categorising these into visualisation methods (making a further distinction between ‘mapping’ and ‘modelling’), narrative-based approaches and indicator-based approaches.³⁶

A handful of such tools have since been applied in USAID-funded programmes in the WASH sector, specifically. The 2016-21 USAID-funded Sustainable WASH Systems Learning Partnership (SWS) is the main example, which is testing systems approaches in different WASH subsectors (rural and small town) across Uganda, Ethiopia, Kenya and Cambodia.^{vi} SWS applies a range of approaches to monitor specific country activities and compiles data at programme level.

Among these, innovative supplementary approaches that could help sharpen the focus on system interactions and relationships include **organisational network analysis** and **factor mapping**. The former is being applied ‘to identify patterns of interaction in the network, well-connected and influential (as defined by their network) actors and groups, and network gaps or opportunities’.³⁷ The latter is a systems tool for identifying the ‘factors that influence systems, and mapping all possible influences that exist between these factors in systems diagrams’.³⁸ So far, baseline analyses have been conducted in the project’s local areas and, in the case of network analysis, there is an intention to repeat analyses and measure changes (interview data). This itself opens up important questions about what key variables will be measured and how changes, e.g. in ‘network strength’, will be assessed.^{vii} Moreover, as USAID notes, ‘Mapping techniques that capture systems change over time can be powerful tools *when linked to adaptive project implementation modalities*’.³⁹ This requires simpler methodologies and easily intelligible summaries of what results mean. The SWS partners have reportedly worked to reduce the complexity of network analyses, which will be essential if they are to be useful for monitoring and evaluation in rapid cycles that can inform adaptive programming.

3.3 Incentive risk 3: Overlooking emergence

A straightforward response adopted by some building block frameworks, is to allow for the inclusion of new blocks or indicators. For example, WaterAid’s Sector Strengthening Programme Design Toolkit encourages collective identification of ‘critical sector processes and functions required for service sustainability’ *before* these processes are grouped under seven pre-defined building block headings, while also encouraging discussion of ‘processes that don’t fall neatly under a building block’.⁴⁰ This flexibility presents obvious tensions with the central need for

^v See www.usaid.gov/GlobalDevLab/MERLIN/SPACES-MERL.

^{vi} See www.globalwaters.org/SWS.

^{vii} There are accepted metrics within network analysis scholarship for variables such as cohesion, density and the strength of ties within a network, though they require expertise for proper analysis.

monitoring, as opposed to one-off assessments or periodic reviews, to use consistent metrics that enable analysis over time (see box, Introduction). As emergent behaviour is invariably unique to a particular system, it is also unlikely that indicators that capture emergent behaviour can be compared between national or subnational jurisdictions. These tensions apply to all monitoring frameworks, not just those based on building blocks. In principle they can be managed by maintaining a consistent core of indicators within the overall set, while allowing some new indicators to be incorporated to track emergent properties (and other redundant indicators, potentially, to be dropped).

A more challenging question is how to ‘spot’ emergent properties if they are not being measured in the first place – there is a lack of robust methodologies to do this across all sectors.⁴¹ One response is to provide a clear structure and set of prompts for stakeholders to identify emergent properties. For example, the SWS project is also using **outcome mapping** as another approach to monitor systems change, with space to record unexpected changes, as well as changes stakeholders would ‘expect’, ‘like’ and ‘love’ to see (interview data).

A similar approach could be adopted by donors in developing programme **project logframes and results frameworks**, by including indicators that are more loosely defined. DFID’s 2016-20 Urban Sanitation Policy Programme, for example, implemented by WSUP, includes as an output indicator within the logframe: the ‘number of significant policy changes catalysed as a result of this programme’. A degree of structure is provided: ‘significant’ and ‘catalysed’ are defined up front, and claims are to be assessed as part of independent evaluations of the programme.⁴² In principle, this does allow system strengthening activities and outcomes that are not predicted at the start of a programme to be monitored, given some value, and therefore incentivised. At the same time, this may bring its own risks; for example, the lack of specificity could allow partners to push inappropriate or unneeded policy change – which could itself encourage isomorphic mimicry. To mitigate this, partners could be encouraged to justify the inclusion of any emergent policy change, captured in this way, against an overall theory of change.

Going a step further than including a single category or indicator for capturing emergent change are methodologies that work largely retrospectively. Usually, these seek to identify significant outcomes, trace back to establish a narrative for how the change occurred, and then identify the role of the intervention within that process. **Most significant change (MSC)**, for example, is a methodology that moves away from the use of indicators altogether. MSC encourages the identification of broad ‘domains of change’, the collection of narrative stories about changes, and prioritisation of these by panels of stakeholders (quantitative data can also be collated to evidence each story).⁴³ **Outcome harvesting** works on similar principles, though relevant outcomes can be identified as and when they emerge, with teams working backwards to establish how the intervention has contributed, directly or indirectly. Experience applying such techniques in WASH reveals the need for careful design and implementation at different stages, to maintain integrity.⁴⁴

The incentives arising for recipients from monitoring in this way could be less immediate, in that during implementation it may be less clear what will ultimately be a significant change or harvested as an outcome, and therefore what actions should be prioritised. Donors, too, may face challenges in justifying funding decisions if there are no pre-determined expected outcomes or impacts. The techniques therefore seem likely to serve as complements, rather than replacements, for conventional monitoring, which forecasts results, change pathways and indicators up front. Cumulatively, however, the techniques could encourage an ethos of experimentation, and help to build a collective, evidence-based understanding of how change has happened, and a programme's contribution to it.

3.4 Incentive risk 4: Misplaced effort

If the retrospective techniques just described will not be sufficient on their own, it remains important to select the optimal indicator set, in advance. Understanding what the most important metrics are to focus on when trying to monitor system change avoids misplaced effort in monitoring unimportant issues, and, more important, misdirecting reform effort on the basis of misplaced monitoring.

To date, most building block frameworks have addressed this by aiming to reduce building blocks and indicators to a focused list based on norms and consensus.⁴⁵ Scoring systems are also often normative – most of the frameworks mentioned in Section 2 involve some kind of scoring or categorisation of the strength of a building block or underlying indicators. Evidence also plays a part: assessment of existing studies and reports has reportedly informed indicator selection in some cases.^{viii} Yet, the general lack of evidence about what institutional forms have supported what outcomes in WASH, through what pathways, has made this difficult.

A first broad group of responses to improve the empirical basis for selecting indicators involve ex-ante analysis. **Quantitative comparative analysis** could use existing datasets, for example, comparing GLAAS data on changes in WASH systems with JMP data on outcomes (initially access but potentially also services as further SDG indicator data are collected). More **qualitative comparative analysis** is also possible, such as the World Bank's WASH Poverty Diagnostics⁴⁶ and the CSO synthesis reports.⁴⁷ However, these did not seek to trace the relationship between specific interventions in systems and system-level outcomes.

More in line with this latter idea, GIZ recently commissioned an independent assessment of various urban water sector reform processes that unfolded from 2005 to 2015 in five countries in Africa. The study confirms the importance of general themes such as corporate governance for utilities, enabling policy and regulatory frameworks, leadership, and functional information systems (with external verification and auditing of reported data an important sub-component). It also notes, however, that this does not equate to a blueprint of certain institutional models (linking back to Risk 1: Isomorphic mimicry). For example, separate regulatory agencies were found

^{viii} The second CSO report, for example, cites 15 country and regional studies to explain the choice of indicators introduced over and above the first round. (AMCOW, 2011)

to have been useful in some contexts, but their functions could also be effectively fulfilled through other arrangements.⁴⁸

Innovative research methods can also be used to improve robustness when looking across existing case study evidence, to identify important cause-effect relationships in certain generic types of WASH system. For example, ‘fuzzy-set qualitative comparative analysis’ has been applied ‘to identify the combinations of causal conditions associated with successful long-term functioning’ of various rural water supply setups – including piped networks, self-supply and community-managed water points. Drawing on more than 8,000 articles, narrowed to 20 case studies, the research found good financial management and user participation in project decisions were important across all setups.⁴⁹ Again, these types of analysis could provide a starting point for deciding what are the most important areas or building blocks to monitor in equivalent WASH setups or systems.

These forms of more robust comparative analysis can help to refine the most important binding constraints across generic types of WASH sector or subsector systems, including as a basis for reducing the number of themes or building blocks. However, this must simultaneously leave space for stakeholders to tailor their monitoring, and interventions, in response to variations in local context. Otherwise, addressing Risk 4 could exacerbate Risk 1, i.e. promote best practice blueprints that incentivise internationally accepted ‘form’, not locally relevant ‘function’. It is therefore equally important to empower those embedded within the system to prioritise and select some building blocks, and certainly the detail of specific metrics to be monitored, based on locally identified problems. **Problem-driven analysis** has an important role here. The Hilton Foundation is reportedly moving in this direction, in agreeing system change objectives and key indicators with its grantees. Grant recipients are asked to identify just two to three key dysfunctions that prevent the desired outcome from being achieved; the underlying causes, with evidence; and which of these they can address. This selection then provides a basis for monitoring (interview data).

This requires donors to relinquish some control over what will be prioritised and monitored to recipients – which some donors may struggle with. A final option is therefore for the donor to specify one or more desired result, but to do so some way down the causal chain, leaving the recipient to work out how to get there. The incentive could be sharpened by attaching funding to the monitoring and achievement of those results: the principle behind many variants of **results-based funding** (including both results-based aid – contracting a government, and results-based financing – contracting an NGO).⁵⁰

In WASH, there is mounting experience with result-based modalities that specify service outputs and outcomes as the payment results, for example, DFID’s WASH Results Programme or the Global Partnership for Output Based Aid. There is debate about whether results-based funding can incentivise system strengthening in practice. Much depends on what results are specified, who is responsible for

achieving them, and whether they can be achieved by bypassing systems more cheaply and effectively than by strengthening them.⁵¹

Given the doubts, there is increasing interest in attaching financing to changes in the system, rather than the outputs or outcomes it produces. For example, GIZ and WSUP have separately investigated similar concepts for results-based funding, in which utilities would be able to access sequenced financing if they meet thresholds for sound management and governance.⁵² Similar approaches are being initiated in Indonesia (discussed in the next section). By themselves, these approaches don't solve the problem of how to select which properties in the system to monitor, as the basis for payment. Indeed, as GIZ notes, the concept is predicated on 'identifying the key leading indicators that reveal the conditions for sound governance substance (rather than form), and their opposite (compromised governance)'.⁵³ Nonetheless, over time, the approach could accumulate useful and rigorous datasets on which types of system strengthening intervention have been most effective for improving performance. This could then provide a basis for focusing subsequent monitoring efforts, results for payment, and corresponding system strengthening activities.

4 Monitoring sector systems in practice: Country case studies

The preceding section focused on various innovations that have predominantly originated from donors and international agencies. However, in practice, approaches to assess and monitor systems tend to co-evolve among the actors concerned with a given sector, as a product of collaboration and negotiation. Arguably, the most important group of actors are national and subnational governments, who are ultimate custodians of the sector systems. This section considers the co-evolution of approaches between government, donors and others in three sectors – WASH and health in Ethiopia, and WASH in Indonesia – before comparing how the four key incentive risks have been addressed across the three contexts.

4.1 Indonesia's WASH sector

Indonesia's WASH sector features several important innovations relevant to system strengthening. Given the focus on system monitoring, however, this case study focuses on Indonesia's comparatively rich experience with results-based funding in the urban water and sanitation subsectors,^{ix} and how this has progressively acquired greater focus on systems-related elements. The experience demonstrates the challenges of finding robust indicators that satisfy both as a financial tool (as a trigger for payment) and as a governance tool (as an incentive for system strengthening), as well as some interesting potential solutions.

Indonesia is of interest because while the main results-based funding innovations have generally come from donors, they have been enthusiastically adopted by the

^{ix} While there has been more experimentation on the urban side, one of the longest established example is in the rural subsector. The Water Supply and Community Based Sanitation Program, PAMSIMAS, has been running since 2008 in rural and peri-urban areas, with support of the World Bank.

national government, as a way to encourage local governments to strengthen their WASH systems. Indonesia's wider political context is marked by rapid and far-reaching decentralisation since 2001. This has provided significant authority and autonomy to the second tier of subnational government, the *kabupaten* (regencies) and *kota* (cities) (hereafter, both are referred to as districts). In practice, however, many districts fail or are unable to take on these responsibilities, for example, in following up investments in water or wastewater treatment plants made by the national level, with piped water or sewerage networks. Subnational expenditure on WASH is low, relative to requirements, and the central government has limited mainstream instruments with which to encourage more, and more considered, spending on WASH by the districts. Central government also has limited ability to monitor how those transfers that can be earmarked to particular purposes (a small share the total) are actually spent, and what is achieved with these funds.⁵⁴

A first example, from urban water supply, is the *hibah* scheme.^x This is an output-based aid modality, first piloted in 2010 with support from the Australian Government, and more recently mainstreamed by the Government of Indonesia. The scheme operates by reimbursing district governments an agreed amount for verified water supply connections, targeted at low-income households. As payments are made on the basis of connections, the scheme monitors system outputs, rather than system strength. However, the original objective was more systems-oriented: to shift the incentives for local governments to invest in their water utilities (interview data). A recent evaluation found that the *hibah* has succeeded in encouraging local government to allocate its own budgets to WASH, and that it may have had other incidental system strengthening effects, such as improving accountability for funds and the efficiency of intergovernmental transfers.⁵⁵

However, progress in extending connections is a crude and imperfect proxy for performance, and in some cases the focus has led to unintended consequences and even declining service outcomes. Eligibility for financial assistance under the *hibah* schemes is restricted to districts with stronger WASH systems, judged based on various technical and performance metrics collected by a central government auditing agency⁵⁶. Poorer performing utilities, with high non-revenue water, have nonetheless been able to take part. Where they expand connections without first enhancing bulk water supply, it has resulted in a drop in service levels for those already connected (interview data).

The output-based emphasis of the original *hibah* schemes reflects a wider prioritisation of access within Indonesia's WASH monitoring. Mainstream monitoring mechanisms in the WASH sector, such as the national WASH MIS (NAWASIS), still focus predominantly on access and infrastructure development. However, several interviewees reflected on the need to move towards a more 'system-oriented' approach to monitoring that can track strengthened governance. In line with this, the Governments of Indonesia and Australia, as well as the World Bank, are planning to move to 'performance-based grants' for districts, which would be released based on

^x There are also variants of the *hibah* scheme for sanitation, for example, reimbursing local government for increasing sewerage connections.

assessment of elements of the WASH system, rather than service outputs. In the new National Urban Water Supply Project (NUWSP), prepared under the Government's National Urban Water Supply (NUWAS) framework in collaboration with the World Bank, the payment indicators are restricted to two measures of operational performance that can be assessed straightforwardly and quantitatively: reduction in non-revenue water; and improvement in energy consumption. To avoid the unintended consequences mentioned above, these metrics are to be cross-checked against service quality (e.g. pressure) and number of connections.⁵⁷

The programme supported by the Australian Government, however, plans to supplement these operational performance payment indicators with ones that look more directly at aspects of system strength, such as governance. These include existence and frequency of reviews by utility oversight boards; and evidence of a water utility business plan being approved by the head of local government – as well as financial indicators, such as operating ratio and billing efficiency (interview data).

A second example comes from urban sanitation. Here, improvements in district sanitation governance are tracked to determine eligibility for financing from central government. One example of such 'readiness criteria' is the presence of a city sanitation strategy. Such strategies have been promoted in successive urban sanitation programmes for over a decade. On one level, this has been successful – more than 450 districts now have such a strategy and implementation plan, and the approach is fully embedded within Government of Indonesia processes. However, there are persistent challenges with the quality and implementation of the strategies, especially in the absence of strong political commitment from the elected district head, and clear legal status for the strategies themselves (interview data). Case study research suggests that, at least in some instances, the strategies are prepared solely to access funding and never implemented.⁵⁸ This implies that the risks remain, of both isomorphic mimicry (Risk 1) and of overlooking key system-level issues (such as political commitment at different levels – Risk 2).

Among various responses, the World Bank and UNICEF have been working on an assessment model for tracking local government readiness for urban sanitation investment, at the invitation of the Indonesian Government. The model is based on a building block approach, drawing on the SDA/CSO service delivery pathway for its overall framing, with building blocks related to enabling, developing and sustaining services. For more detailed indicators, it builds on those developed in a USAID-funded project working with 32 local governments across Indonesia.

The framework includes 18 indicators, ranging from the type of institution managing domestic wastewater,^{xi} to the ratio of tariff to operating cost. Scoring thresholds are quantitative wherever possible, and where they are qualitative they are highly specific. This could imply the same risk of promoting isomorphic mimicry, though the indicator assessment criteria are closely tailored to Indonesia's existing institutional context (interview data). For example, for the type of institution managing domestic

^{xi} The term 'wastewater' is frequently used in Indonesia to distinguish both water-borne and onsite sanitation from solid waste management, which is also considered part of sanitation ('sanitasi').

wastewater, progressively higher scores are obtained for having one of eight, legally recognised, institutional forms, each with a greater degree of autonomy from the local government. While a draft has been trialled in 16 cities, the framework has not yet been operationalised as a basis for targeting and accelerating investments. However, there are suggestions that it could be backed by a range of funding mechanisms, combining standard direct grants with incentive grants that would be released based on indicators such as those mentioned, as well as indicators of service performance. As with the building block approaches discussed in Section 2, there is an intention to use the framework for ongoing monitoring of progress as well as an initial assessment of investment readiness (interview data).

In summary, the various approaches across Indonesia's WASH subsectors indicate the enthusiasm of both the Government of Indonesia and its development partners to find ways of incentivising WASH system strengthening on the part of the districts, and to move beyond output-oriented forms of results-based funding to do so. The search is underway to find a set of measures that can better satisfy as both financial and governance tools. It should be noted, however, that neither the performance-based grants for water supply nor the more sophisticated readiness criteria in urban sanitation have been implemented at scale or for any length of time. It remains to be seen whether these can counteract the prevailing emphasis on infrastructure development in both political priorities and budgeting in Indonesia, which continues to exert a strong influence on district governments.

4.2 Ethiopia's WASH sector

Following significant achievements in the last decade or so, Ethiopia is often cited as a key example of system strengthening in the WASH sector. The 2005 Universal Access Plan, and its 2011 successor, provided a high-level statement of ambition and were key to galvanise financial and political support to the sector, including by making a link to poverty reduction. This laid a path for the One WASH National Programme (OWNP), a sector-wide approach (SWAp), which was launched in 2013. Together with several other reforms and policy developments, this has resulted in a marked consolidation of the WASH 'system'. Preparations for the second phase of the OWNP are well underway, updating and aligning it with the latest national development strategy (the Second Growth and Transformation Plan, GTP II).⁵⁹

As we discuss below, progress in system strengthening has so far been achieved and sustained without any robust routine monitoring of system strength. In terms of what has driven and shaped the effort to date, interviews pointed to the importance of politically valuable 'results', such as announcement of the achievement of the MDG target on water supply in 2015. To a lesser extent, reduction in open defecation has also been important in giving confidence that the sector is moving in the right direction.

This is not to say there is no consideration of system-related issues. More-or-less annual review processes tied to the OWNP cycle have played a role in helping all parties – government, donors and others – to identify general trends in system

strengthening. These include a joint technical review (JTR) conducted by government and partners and a multi-stakeholder forum (MSF) for discussion and agreeing actions. In principle, these also allow specific system-level weaknesses to be identified and responses jointly developed. Since this is done on a rolling basis, it might address the challenge of spotting and addressing emergent challenges (Risk 3). However, there are difficulties in following up on agreed actions in the absence of a formal accountability mechanism – in effect, this undermines the prospect of addressing any identified challenge, whether old or new (interview data).⁶⁰ The MSF and JTR are also largely restricted to national level – instituting equivalent, and more regular, processes at regional and woreda level is a recognised priority for OOWNP Phase II.⁶¹

The review mechanisms associated with the Consolidated WASH Account (CWA) appear to be more effective in this regard.^{xii} While the CWA involves only a subset of the sector's development partners, woredas and expenditure, the system to agree follow-up actions and check these each year is reportedly stronger. CWA data also dominates the second of only two OOWNP-wide reports produced to date, in 2015/16 and 2016/17 (interview data). This seems to suggest that the joint commitment of funding may be helping to strengthen reporting and mutual accountability.

Rather than an explicit 'building blocks' approach, the OOWNP documents have provided the main framing and structure for how the system is conceptualised and reviewed, if not monitored in the routine sense. For example, the OOWNP review reports are organised according to the four main pillars of the OOWNP – rural WASH, urban WASH, institutional WASH, and programme management and capacity – with additional chapters on emergency WASH and sector investments. There is consideration of more process or systems-oriented elements, though most of the 17 key performance indicators (KPIs) of the OOWNP are for service outcomes (e.g. access, performance) and impacts (e.g. health, education, time savings). Among the OOWNP KPIs there are a handful of indicators relating to institutional issues, for example, the presence of active community-based management structures, and levels of cost recovery. However, in both cases, limited data on these were available in the two OOWNP reports released to date.⁶²

This reflects a wider focus on monitoring what the system delivers, rather than its strength, in Ethiopia's WASH sector. It also presents some uncertainties as to how effectively the more than 70 indicators in the OOWNP Phase II results framework will be monitored and used.

^{xii} A mechanism to pool contributing partner funds and route these through the Ministry of Finance and Economic Cooperation (MoFEC) system. Currently financed by the UK and Finnish governments, the African Development Bank, UNICEF and the World Bank, the CWA covers approximately a third of woredas – several donors continue to use their own financing and implementation modalities (GoE, 2018a).

The latest programme plan presents many more indicators that appear to be aimed more at monitoring system strength, though they are designated at output level.⁶³ These include:

- Enabling environment and good governance:
 - Number and % of WASH committees with legal status
 - Number and % of quarterly WASH review meetings held by level
- Efficient use of resources
 - % reduction in per capita investment costs in town, cities and rural areas
 - Number and type of pilot/demonstration activities scaled up
- Strengthened capacity of WASH sector actors for achieving and sustaining results
 - % of post-construction support units established and functioning
 - Number and % of water quality laboratories supported and operational

As with some of the examples of governance indicators in Indonesia mentioned above, the emphasis on counting certain institutional forms or processes leaves a residual risk of isomorphic mimicry (Risk 1) – for example, meetings being held but not resulting in any meaningful decisions. The large number of indicators may also not just be overly ambitious but lead to a loss of focus (Risk 4).

The wider systems for routine monitoring are better suited to tracking service outputs and outcomes than system strength. They do, however, provide some lessons for future efforts to enhance systems-oriented monitoring. Foremost is the challenge of ensuring coherence across the many sectors participating in WASH in Ethiopia, as elsewhere. Data is sourced from the health and education management information systems, household surveys and other monitoring efforts led by the Ministry of Water, Irrigation and Electricity, such as the second round of the National WASH Inventory (NWI2), which will focus on water supply schemes.^{xiii} Where data are successfully obtained, they are also used more for upward reporting and periodic planning (e.g. the five-year cycle), than for regular performance monitoring or course correction.⁶⁴

Interviewees identified limited cases of externally derived building block frameworks being used at national level, though a workshop-based analysis was undertaken using the WASH BAT tool as part of the final review of OWN Phase I and preparation for Phase II.⁶⁵ Examples of progress monitoring using building blocks were not identified at national level in the WASH sector – the closest being IRC and WaterAid, both of whom are attempting to use their respective building block frameworks to inform their system strengthening activities at woreda level.⁶⁶ Also of interest, albeit as yet at a small scale, are the activities of the SWS programme in Ethiopia. A baseline ‘local systems analysis’ in two small towns includes

^{xiii} See Annex 3 ‘New initiatives’ in GoE (2018) for a full list of monitoring and review initiatives.

organisational network analysis and factor mapping, mentioned in Section 3.2.⁶⁷ This was a rare example identified in the case study of an attempt to consider system level dynamics (Risk 2). However, it remains to be seen how the baseline will be updated for progress monitoring purposes.

To conclude, effort to strengthen Ethiopia's WASH sector on the part of various ministries and donors has been sustained, so far, in the absence of a methodical approach to system monitoring. The key question is whether that effort could be more effective, with better and more regular data on the system strength. An established cycle of sector reviews do consider institutional elements. However, the focus of these changes over time, and there is a lack of follow-up (there seems to be slightly more accountability where there is money on the table, in the CWA). Consideration of systems-related elements is minimal in the current SWAp KPIs and the relevant data has been difficult to obtain. There are plans to introduce more indicators of this kind for phase 2. Initial indications are that these may be quite superficial in nature, measuring institutional form rather than function. At district level, there are some interesting innovations supported by NGOs, including some of the methodologies identified in Section 3, to monitor system-level interactions.

4.3 Ethiopia's health sector

Ethiopia's experience of monitoring its health system shows several parallels with the WASH sector. Monitoring of system strength is again done more through successive and evolving five-year plans, around which there are well-established joint-sector review and reporting processes, than through routine monitoring systems. As in WASH, the Government can exert a firm hand over the activities of development partners, coordinating reporting under the principle of 'one plan, one budget, one report'.⁶⁸ Progress at the macro-level in terms of health outcomes has also helped reinforce donor commitment, for example, around key health-related MDGs, notably under-five mortality and maternal mortality.

There are some differences, however. Arguably, the ability to ensure all development partners submit to the 'one report' is greater than in WASH, which may be related to having a single ministry in charge of the sector. The SWAp in the health sector, and related monitoring systems, have also had longer to evolve – initiated in 1997 with the first five-year Health Sector Development Plan, the approach has evolved through three further cycles, currently culminating in the 2015-20 Health Sector Transformation Plan (HSTP).

Although they are more of a review mechanism than a monitoring system, the well-established cycle of annual performance review reports, meetings, and joint government and partner missions provide an update on trends in service delivery, but also examine progress in system strengthening. The 2016/17 report, for example, provides updates on the four 'transformation agendas' of the HSTP: equity and quality of health care; woreda transformation; the information revolution; and a compassionate, respectful and caring health workforce.⁶⁹

The successive five-year plans themselves also provide a thematic structure and inform which issues are looked at, rather than any wider or longer-term conceptualisation of the health system using building blocks or another framing. More substantive review processes are undertaken to inform the development of each new five-year strategy, and each strategy is also reviewed in some depth at mid-term. However, building blocks have been applied in planning exercises; the WHO health system building blocks provided a framework for costing of HSTP, and reportedly informed both USAID and the Bill and Melinda Gates Foundation in their approach to supporting the health sector in Ethiopia. UNICEF also plans to integrate an online version of its health bottleneck analysis tool into the health management information system, which is currently being overhauled (interview data). Examples of progress monitoring using building blocks were, again, not found at national level.

Looking in more depth at the indicators in the current HSTP, just over a third of the 176 specified are related to health system performance. Some do attempt to measure 'function' rather than simply 'form' (Risk 1), and system-wide aspects such as governance and learning (Risk 2). As in the case of some of the sanitation governance indicators proposed in Indonesia, some are prescriptive but use locally relevant standards and criteria. This may improve the prospect that they will encourage relevant 'forms', even if they don't always look at function. Indicators include:

- Indicators related to an improved regulatory system, for example:
 - 133 'Number of licensed, ethical and competent health professionals'
 - 136 'Number of healthcare facilities meeting healthcare facility standards'
- Indicators of improved research and evidence for decision making, for example:
 - 157 'Proportion of health institutions that met minimum information use standards/criteria (regular performance review with plan vs achievements, root cause analysis, charts/figures display, action plans, shares responsibility and track implementation of endorsed plan)'
 - 158 'Proportion of synthesised evidence-based information utilised for decision making'

Other indicators also measure dimensions related to community engagement, equity, efficiency and effectiveness.⁷⁰

However, according to interviewees, the service output and outcome indicators, which dominate the list, are also the focus for decision makers. A possible exception is the World Bank-managed SDG Performance Fund,^{xiv} a Programme for Results (results-based aid modality) in which disbursements are linked to several indicators.

^{xiv} Formerly the MDG Performance Fund. The MDG Performance Fund includes 8 Disbursement-Linked Indicators (DLIs) and 12 sub-indicators. The SDG Performance Fund restructures, changes and adds DLIs for a total of 15 with 20 sub-indicators.

These mainly relate to service outcomes, but include some on system aspects, for example, ‘number of health centres reporting HMIS data on time’ and ‘undertaking surveys and disclosing results with actions to address weaknesses’.

More generally, monitoring in Ethiopia’s health sector raises two important broader points around the political economy, which would be just as relevant for more systems-oriented monitoring. Firstly, although policy attention is focused more on indicators of service outputs and outcomes, the data is not necessarily used for routine decision making. HSTP indicators are intended to draw on a range of sources, including the Ethiopian Health Management Information System (HMIS), periodic surveys and administrative reports. However, discrepancies have emerged between the different data sources and efforts to rationalise are underway.⁷¹ Interviews suggested that data from household surveys are generally seen as more trustworthy, and that the discrepancies had a corrosive effect on stakeholders’ ability to trust, and therefore use, the HMIS data.

Secondly, political ideology can play an important role in decisions about system reform, sometimes trumping evidence. For example, the belief that health insurance schemes will be able to address a substantial share of the gap in finance for health has been argued to be ‘based on a leap of faith – rather than strong evidence’, because it resonates with core socio-political ideals like ‘the importance of everyone contributing what they can and the resistance to anyone receiving support without contributing something in return’.⁷²

In both cases, this implies that even if indicators of system strength are better designed and used, stakeholders need to be wary of the politics around monitoring; data of any kind needs to be verified if it is to be trusted, and ideology can often play as great a role in decisions as evidence.

In summary, the monitoring framework for the current plan, the HSTP, includes some interesting examples of system-related indicators. However, upward reporting of service output and outcome data reportedly dominates officials’ attention, even if use of this data for effective decision making is undermined by discrepancies between different sources and the strong role of ideology. As in Ethiopia’s WASH sector, the apparent lack of effective routine monitoring of system strength has not prevented donors from sustaining their commitment to the sector, over a longer timeframe. The review mechanisms and periodic revision of plans have again played a role, offering confidence that the system strengthening journey is generally progressing (and helping identify, over time, where it needs to be adjusted), but also face the same problem of not constituting a framework for structured monitoring over multiple years and programme cycles.

4.4 Addressing the incentive risks in country-level system monitoring

How, then, do the mechanisms described in each sector address the key risks by effectively incentivising a focus on function, system-level interactions, and emergent and salient issues?

Incentive risk 1: Isomorphic mimicry

Examples of indicators that measure how an institution is functioning, as opposed to how it looks, are somewhat in evidence in Ethiopia's health sector. They are also proposed in some cases for future variations of results-based funding in Indonesia's WASH sector. In both cases, however, interviews would suggest they hold less political importance than indicators relating to service outputs and outcomes. The best examples would appear to assess whether an organisation is doing what it is meant to be in some objectively verifiable way – publishing minutes or approving plans, for example. Some examples specify the form that institutions should take in much greater, locally relevant detail. While this is unlikely always to show how institutions are really performing, it does mitigate the risk of isomorphic mimicry by increasing the likelihood that the desired form is appropriate for the local context. In Ethiopia's WASH sector, meanwhile, there are limited examples in the current SWAp, and those proposed for the second phase seem mainly to assess institutional form.

Incentive risk 2: Siloes not system

Few convincing examples of indicators or monitoring approaches to tackle this aspect were identified across the cases studies. Network analysis and factor mapping have been used to assess system interrelationships as part of the USAID-funded Sustainable WASH Systems Learning Partnership activities in Ethiopia's WASH sector. It is too early to assess how viable or effective these approaches will be for periodic monitoring as opposed to initial diagnostics. In Ethiopia's health sector, some interesting indicators are included in the HSTP which seek to measure system-level phenomena, for example, around the use of research and evidence. Indicators proposed for measuring local governance and commitment in Indonesia's urban water sector could offer similar inspiration, though they are as-yet untested.

Incentive risk 3: Overlooking emergence

In both Ethiopia's WASH and health sectors, the periodic, five-year planning cycle and more frequent sector reviews are, at one level, well suited to identifying emergent issues. New priorities for system strengthening can be set with every iteration. However, this also makes it harder to track the same issues over the long term, which would help ensure institutional memory and allow for more systematic prioritisation of resources and reform effort. Additionally, the lack of accountability for follow-up, raised as an issue for the WASH sector review process at least, undermines their effectiveness for remedying both emergent and long-standing

failures within the system. No convincing responses to this risk were identified in the Indonesia case study.

Incentive risk 4: Misplaced effort

The monitoring frameworks for the HSTP in Ethiopia's health sector features 176 indicators, with the result that there is less scope to pay attention to some of the more interesting, systems-oriented indicators. A more focused set has been adopted for the results-based SDG Performance Fund, suggesting some effort has been made to identify the most important proxy indicators for system strength. In WASH, the smaller set of KPIs for the OWNPN looks to be expanded significantly for Phase II. While systems-related indicators are more in evidence, it seems likely that this expansion will amplify existing challenges of data collection. Limited decision making based on even output-level monitoring data does not offer much confidence that the new indicators will be used. Indonesia's evolving approach to selecting indicators for results-based funding in WASH has yielded more focused indicator sets, since excessively numerous indicators would make verification and payment much more complicated. So far, the relevant indicators appear to have been selected based on theory and discussion. However, given the diversity of existing and planned approaches, the opportunity arises for a more empirical approach: comparing data across different mechanisms and between different districts, to identify the most salient proxy indicators for system strength.

5 Ways forward: Conclusions and recommendations

5.1 Conclusions

In terms of *how WASH system strengthening is currently being monitored* (research question one) most systems-oriented monitoring frameworks promoted by donors and NGOs are based on building blocks. While there is some emerging consensus in the international community, for example, around the Sanitation and Water for All building blocks, several new frameworks have been developed in the last two years, particularly at the subnational level. However, despite continued enthusiasm for building block frameworks, to date, their use for monitoring, as opposed to diagnostics, has been limited. Some NGOs, for example, Water for People, as well as WaterAid Ethiopia, have conducted more than one round of assessment at the subnational level. At the national/international level, GLAAS continues to monitor systems-related elements. This has built a multi-year dataset, but it did not appear, from interviews, that this was being used by donors to monitor the system strengthening impact of their investments. At the same time, GLAAS may play a larger role in helping country governments, who report against the indicators, to consider their WASH systems' strength in a regular, structured format. In the country case studies, building block frameworks did not appear to be used at sector level by either the governments or donors for monitoring – though in Ethiopia's WASH and health sectors, frameworks have been used as diagnostic tools to support planning and budgeting.

There is also limited evidence to date that analyses of sector strength, whether as one-off diagnoses or ongoing monitoring, have a major influence on strategic policy or investment decisions by donors. DGIS's sustainability checks, which include a scorecard assessment grounded in FIETS categories^{xv}, is a possible exception. The accompanying sustainability clause draws a relatively clear line between the assessment and finance. This is not quite the same as results-based funding, where payment is only allocated if something is achieved (rather than a theoretical financial penalty if something fails later). Most results-based financing instruments in WASH identified by this study trigger payment on the basis of output or outcome results, rather than system-level changes, though efforts to explore performance-based funding modalities, including in Indonesia, suggest that this is starting to change.

This is not to say that donors are ignoring systems altogether – we have highlighted numerous examples of investments to understand and strengthen WASH sector systems. However, it is not clear that *monitoring* of systems, as yet, plays a major role in shaping incentives for investment within specific programmes or portfolios. Where system-related indicators are monitored, they appear to be, at most, a supplement or addition; the main indicators that matter remain access or coverage. The SDG indicators may move the conversation beyond purely household-level infrastructure, but the results that matter most still tend to be those that can be counted in terms of beneficiaries.

For governments, meanwhile, our conclusions are limited by scope – this research considered only two countries. Indonesia, as a middle-income, highly decentralised country, is embracing opportunities to use finance to incentivise subnational government and utilities according to different measures. Together with donors, the focus has to date been more on rewarding improvements in access (e.g. the water and sanitation *hibah*) or simplistic measures of institutional strength (e.g. completion of city sanitation plans as a readiness criterion for investment). Yet there are signs that the Government and development partners are identifying and starting to apply more sophisticated metrics of sector strength. Ethiopia, meanwhile, has anchored its assessment, review and monitoring of the strength of both the health and WASH sectors in its own sectoral strategies, rather than any external building block framework. At one level this represents a locally owned, adaptive and inclusive process. At another, the review processes appear to have a constantly evolving focus. The current health SWAp plan, and the next one for WASH, include some examples of systems-oriented indicators. These appear more carefully thought through in health. However, to date, the focus of attention in both sectors is on service output and outcome indicators. Reviews, particularly in WASH, have also offered ample opportunity for diagnosing problems, but fewer hard incentives for specific actors to own or follow up on addressing these problems.

Noting that monitoring of systems is not currently a primary factor shaping donor, government or implementing partner decisions does not reduce the need to answer

^{xv} The FIETS approach outlines five areas of sustainability: financial, institutional, environmental, technical and social.

the other research questions: What sort of behaviours/activities/programmes does system monitoring incentivise? To what extent does this align with what we know about the complex and dynamic nature of systems change? What innovations in monitoring complex systems could be usefully applied to maintain a focus on the key drivers of sector performance in a dynamic WASH sector? It is early, but not too early, to be asking these questions. Many donors have committed to increasing their investments in system strengthening. If the WASH SDG targets are to be achieved, donors and country governments will need robust approaches to monitoring, to track progress and direct further effort.

In answer to the second and third research questions, we have identified four key issues that building blocks, and alternative and adapted approaches to systems change, need to address if they are to incentivise constructive behaviour from donors and governments in support of complex and dynamic WASH systems. Because actual examples of coherent WASH system monitoring are so limited, these are framed as risks of incentivising certain behaviour, rather than behaviours that are already widely incentivised. The incentive risks that we identified are in most cases already acknowledged by those advocating building block type frameworks, and those innovating other approaches are grappling with equivalent issues. These examples of innovations from international agencies, as well as some developed at country sector level between government, donors and NGOs, help to address the fourth question.

Before turning to the four incentive risks, it is worth making four preliminary points. Firstly, any critique of building block approaches in this paper does not mean they should be dismissed. There is no single approach that appears more credible and effective, in either health or WASH. It will be more productive to consider how to adapt and combine a range of approaches to framing, understanding and monitoring sector systems, than to claim that approach a or b is best.

Secondly, the wider context for system strengthening is challenging. Despite donor commitment to system strengthening on paper, several interviewees reflected that the political space for systems building is diminishing – with a retreat from budget support, both sectoral and general, in favour of projectised approaches. There is therefore a clear and urgent need for donors and governments to develop credible mechanisms to demonstrate that money spent on system strengthening is money well spent. This is not to say that enhancing system monitoring is the only way to catalyse investments in system strengthening – some system strengthening investments occur, even in the absence of widespread effective monitoring of the systems themselves. However, where this occurs it has often been attributable to specific political economy factors. An example is donor and government commitment to system strengthening in Ethiopia without a structured, routine monitoring effort. Here, confidence in Ethiopia's general capacity to deliver services, for example, through effective decentralisation, has arguably sustained initial commitment to systems building through the SWAps, until outcome results could be demonstrated. In any case, to direct investments in system strengthening, as opposed to just

catalyse them, donors and governments will need more detailed and systematically organised evidence of system-level changes.

Thirdly, the case studies also show that the actual effectiveness of any framework for monitoring systems will be determined as much by the political economy of data use in each jurisdiction, as the careful design of indicators. Examples include discrepancies between sources eroding confidence in data in Ethiopia's health sector; the challenge of finding simple metrics that can create the right incentives in Indonesia; and the greater likelihood of effective accountability where money is on the table in Ethiopia's WASH sector. Although these examples all arise in relation to output or outcome-oriented data, they would also influence how system-oriented monitoring would be undertaken, and the data used.

Finally, the general emphasis on output and outcome-level data observed in both Ethiopia and Indonesia points to a broader challenge: getting stakeholders to take indicators of 'soft' progress, such as governance, learning or leadership, as seriously as they do measures of more tangible or visible progress in increasing access or use.

Considering this context, we draw the following conclusions for each of the four key incentive issues identified:

For the risk of **encouraging isomorphic mimicry, by focusing on form over function**, a priority is to design indicators or building blocks around locally agreed priorities, rather than internationally accepted, yet largely normative blueprints. This will be more challenging where an objective is comparison or aggregation between jurisdictions – which is certainly a concern for many donors. But this is a tension that must be managed, for example, by specifying a core set of key themes but allowing the detail of indicators, and supplementary themes, to be defined through local deliberation. Measures that identify how well institutions really function will also be crucial. Three promising avenues are identified from our review:

- Quantitative performance indicators, for example, levels of budget utilisation, or non-revenue water, as proposed in Indonesia
- Qualitative indicators assessed in a participatory or multi-stakeholder way, or against locally-tailored criteria that describe what an institution or process needs to do (not just look like) – for example, indicators relating to performance of utility board oversight or approvals from local political heads, proposed in Indonesia's urban water sector
- Measures of user experience – for example, the indicators of user confidence and satisfaction proposed to measure health systems

Looking beyond indicators, encouraging closer attention to how institutions need to function in a given system may require an entirely different approach to funding and implementation, as well as monitoring. Methods like problem-driven iterative adaptation or adaptive development have yet to be used extensively in WASH programmes. They would require much more nimble, diverse and long-term

approaches to reform, deploying multiple, smaller value reform experiments – dropping the least promising examples and scaling up the most promising. Related monitoring systems would need to be capable of rapid feedback and potentially monitoring multiple system strengthening efforts simultaneously.

The risk of promoting **siloed responses, by emphasising individual blocks rather than system-level interaction**, is one for which answers are not yet readily apparent. Existing building block frameworks have made tentative steps towards addressing this risk by including blocks and/or indicators relating to ‘system level’ phenomena, such as coordination and learning. However, these are also some of the hardest areas in which to measure function over form. It is relatively easy to identify if there is a sector review or learning mechanism. It is less easy to tell how well learning is happening and how it shapes the sector, for better or worse. Issues such as political commitment and drive also need to permeate throughout a WASH system, but are similarly hard to measure. Here, it is likely the innovations supported by USAID, for example, in the SWS programme, will provide the most relevant lessons in the near term. Methods such as network analysis are intuitively appealing for measuring the strength of relationships and interaction across a system. However, there is an urgent need to demonstrate that they can be used for monitoring rather than research and diagnosis, both at reasonable cost, and in a way in which results can be easily interpreted by non-specialists.

The third risk, of **overlooking emergent issues, by encouraging static rather than dynamic understanding** of the system, is also one in which clear answers are yet to emerge. At one level, Ethiopia’s locally owned WASH and health review mechanisms, by allowing the topics of focus to evolve, are well suited to incorporating emergent issues. However, this flexibility also means issues are not effectively monitored over the long term. Also, there are still few robust methods to identify unanticipated issues in good time (ideally, before they have become a problem). Outcome harvesting and most significant change are possible exceptions but require careful design and management. At a simpler level, it is necessary to encourage a process of reflection at regular intervals to ask, “Is this set of indicators capturing everything we need to know?” Donors could encourage this kind of thinking by putting greater value on non-predicted outcomes in their results frameworks – as DFID has attempted to do by including more loosely specified policy-influencing objectives in WASH logframes – though they will need to specify in some detail what counts as sufficient.

The final risk we highlight is of **incentivising misplaced effort, by failing to focus on the functions within a system that matter most for outcomes**. Here, we identify several approaches that could help. It may be possible to move towards a more evidence-based list of broad priorities for system strengthening, and monitoring, across a range of jurisdictions, using comparative analysis of both qualitative and quantitative data. This could assist with identifying high-level correlations between certain institutional forms and better outcomes. Analysis of country-level datasets would arguably be more useful for this, and less likely to lead to isomorphic mimicry, for example, that which may be assembled in Indonesia’s

urban water sector through the new performance-based grants that use a wide range of performance and governance payment indicators. However, to really avoid encouraging isomorphic mimicry, and yet more focus on best practice blueprints, any effort to identify generally important themes needs to be matched with local, problem-driven approaches. These should encourage the stakeholders working directly within and on the system to identify the most important issues that they can influence, and to tailor metrics and monitoring to these. This could be done by each donor with their recipients (as the Hilton Foundation is doing) or through sector-wide discussion platforms.

5.2 Recommendations

In line with these conclusions, we recommend the following:

WASH donors and implementing partners should continue to deploy building block frameworks, but with greater emphasis on using and improving them for the purposes of recurrent monitoring. This research did not identify any fatal flaws to using building blocks to conceptualise WASH systems, but there are some important risks to be aware of when using them for monitoring. Currently the bigger issue is that they have not been extensively used for ongoing monitoring, especially at country level and below, despite many intentions to do so. Further applications, together with research and evaluation, are therefore needed. The number of different frameworks now available is potentially a strength, if agencies work together to share learning about what is most effective in designing and using them. Within individual countries, however, agencies should consider how to synthesise or align frameworks over time to avoid creating confusion for other stakeholders.

As a starting point towards addressing the risks and misaligned incentives arising from monitoring systems with building blocks, donors and implementing partners should work together on the following priorities:

- **Funding mixed-methods research using datasets that assemble evidence on institutional performance in WASH sectors over time, to identify what matters most for WASH outcomes.** This could be international in scope, for example using UNICEF's WASH BAT analyses, or national, such as Indonesia using results-based funding datasets. Over time, this could provide a stronger basis to select a more focused core set of building blocks or sub-blocks which link to the binding constraints for a WASH sector (or more likely subsector) system. This could also allow for some international comparability, provided blocks are defined with enough constructive ambiguity to allow detailed metrics to be defined at local level.
- **Tailoring the detail of indicators to local context, based on stakeholder consensus, as far as possible.** As implied, even while a core set of building blocks for a given WASH sector type could be progressively refined and reduced, the specific metrics to assess each building block should be developed at a local level, between stakeholders

working directly within the system.

- **Piloting approaches that allow for iterative adaptation and small-scale experiments to identify what is most effective for system strengthening, with accompanying investment in flexible monitoring and rapid-feedback.** To support this effort, WASH stakeholders should establish more links with the growing PDIA and Adaptive Development communities.
- **Designing indicators that capture how institutions function, rather than just what they look like.** This entails quantitative measures of actual performance, but also well-designed qualitative metrics that look at whether key tasks are performed, and perception-based metrics, especially ones that incorporate user perspectives.
- **Further developing and testing methods to capture system-level dynamics such as network strength.** Methodologies like network analysis and factor mapping have so far been mainly used as a diagnostic or research tool. Much needed work is underway to simplify these as tools for routine monitoring and ensure results can be communicated to non-specialists. This effort could, however, be supported by a wider range of donors and partners.
- **Leaving space for unpredicted, emergent outcomes within results frameworks.** These may need to be bounded with guidance and criteria to establish what counts as a significant emergent outcome, to avoid encouraging institutional or policy change for its own sake.
- **Promoting recurrent reviews of a given system-level monitoring framework (whether explicitly divided into building blocks or not) to ensure it is still relevant to the system in question** – including whether it captures emergent issues. Supplementary, retrospective methods such as outcome harvesting and most significant change, can be used to cross-check whether key issues have been overlooked, but may not provide strong, ex-ante incentives to work in certain ways.

Finally, donors and governments need to approach monitoring of systems change as a political, not just technical, endeavour. The above suggestions are predominantly technical responses, around building block selection, indicator design, and monitoring methodologies. The bigger challenge is for development partners and governments to exploit all their available political room to look beyond the narrow results, which reduce everything to infrastructure or anonymous beneficiaries. To help with this, they also need to tell convincing, human-centred stories about how stronger WASH systems make a real difference to people. And they need to continue to foster a culture of learning and evidence-informed decisions based on monitoring data, whether that data relates to the strength of WASH systems themselves, or the services they produce.

Annex 1: List of organisations consulted

Global level

Department for International Development, UK Government
DAI
Directorate-General for International Cooperation, Ministry of Foreign Affairs, Government of the Netherlands
Environmental Incentives
German Corporation for Development Cooperation
IRC WASH
Stanford University
United Nations Children's Fund
United States Agency for International Development, US Government
Water for People
World Bank

Ethiopia – Health case study

University of Gondar
Bill and Melinda Gates Foundation
World Health Organization
Jimma University
United Nations Children's Fund
United States Agency for International Development, US Government

Ethiopia – WASH case study

Department for International Development, UK Government
IRC WASH
Ramboll Finland
United Nations Children's Fund
United States Agency for International Development, US Government
WaterAid Ethiopia
Wollo University

Indonesia – WASH case study

Bappenas, Government of Indonesia
DAI
Department of Foreign Affairs and Trade, Government of Indonesia
Indonesia Australia Partnership for Infrastructure (KIAT)
Royal Haskoning DVH
SNV Netherlands Development Organisation
United Nations Children's Fund
United States Agency for International Development, US Government
World Bank

Annex 2: Timeline of international building block initiatives in WASH

2006 ^{xvi}	A first round of Country Status Overviews under the auspices of the African Ministers' Council on Water (AMCOW), supported by the Water and Sanitation Programme (WSP) of the World Bank, is undertaken for 16 countries in Africa. It introduces a 'sustainability scorecard... designed to provide an assessment, in a quantitative form, of overall sector and sub-sector sustainability, looking beyond simple coverage figures to predict whether access, once provided, will endure.' The scorecard is based on a series of questions 'relating to generally accepted success factors in the institutional and financial aspects of the sector', posed to sector leaders in each country. ⁷³
2008	The GLAAS Pilot Report is released, which combines existing data on WASH finance and coverage with new information gathered through a supplementary survey for countries and external support agencies (development partners). The pilot covers seven countries and is intended as proof of concept for 'a possible global, periodic, comprehensive reporting mechanism to inform policy-making in the sanitation and drinking-water sectors'. ⁷⁴ The country questionnaire includes questions on WASH status and need, policy and institutions, and resources and finance. GLAAS has repeated the survey and reporting on a two-year cycle; the most recent one (2018/19) is currently underway. The questionnaire has evolved, and questions are now grouped under four categories: governance, monitoring, human resources, and finance. The 2016/17 cycle covered 75 countries. ⁷⁵
2009	WaterAid releases a paper, in support of Sanitation and Water for All, on 'aid compacts' build around national plans. This is anchored in a set of five building blocks: policy/strategy, sector coordination, institutional arrangements, sector finance, and performance monitoring. ⁷⁶
2011	Second round of country status overviews, again undertaken for AMCOW with support from WSP, develops and extends the scorecard and applies it to 32 countries in Africa. It frames the 'success factors' as 'building blocks and supplements the first CSO's set 'with additional factors drawn from country and regional analysis'. It also visualises the building blocks according to a 'service delivery pathway', whereby the blocks are arranged according to a logical 'pathway' representing how finance translates to services. ⁷⁷ An accompanying costing tool is used to assess availability of finance to meet national WASH targets. The scorecard analysis has

^{xvi} All years are the year the first report in the relevant series was released; methodologies were often developed, and analysis undertaken, one to two years before publication.

	subsequently been rolled out in equivalent country studies in Asia (where the reports were known as Service Delivery Assessments [SDAs]) ⁷⁸ and in Latin America (where the initiative is known as MAPAS, the acronym in Spanish for Monitoring Country Progress on Drinking Water and Sanitation). ⁷⁹
2012	UNICEF applies the WASH Bottleneck Analysis Tool (BAT). The tool uses both the preceding building block approaches specific to WASH, notably the CSOs, and integrates these into a methodology developed in the health sector to identify bottlenecks to achieving the health and nutrition MDGs, and to facilitate identification of investments needed to remove the bottlenecks (the Marginal Budgeting for Bottlenecks [MBB] tool; Ryan, 2014). The WASH BAT is applied in over 15 countries as an Excel tool from 2012-15.
2016	<p>The Sanitation and Water for All (SWA) global partnership adopts a set of five WASH sector building blocks at its 2016 Meeting of WASH ministers, based on existing frameworks used by partners. The building blocks are accompanied by four ways of working or 'collaborative behaviours' to which SWA partners, including country governments and development partners, had previously committed:⁸⁰ (1) enhancing government leadership of sector planning processes; 2) strengthening and using country systems; 3) using one information and mutual accountability platform; and 4) building sustainable water and sanitation financing strategies.^{xvii}</p> <p>In 2016, the WASH BAT also re-launches as an open-access online tool.^{xviii} The tool is structured according to a set of 'governance functions' developed by UNDP and SIWI⁸¹ and grouped according to the SWA building blocks. The tool is accompanied by a facilitation guide for participatory workshops designed to score governance functions for a selected subsector and jurisdiction;^{xix} identify bottlenecks and underlying causes preventing progress; and define activities, costs and responsibilities to remove bottlenecks. The tool guidance suggests that progress can be reviewed after 1-2 years, and the bottleneck analysis updated accordingly.⁸²</p>

^{xvii} See <http://sanitationandwaterforall.org/priority-areas/the-four-swa-collaborative-behaviours/> for more information. The collaborative behaviours are themselves being monitored by SWA at country level with an intention that they become 'as a tool for ensuring SWA partners hold each other accountable for system strengthening behaviours' (Battle, 2018). Although focused on system strengthening *behaviours*, there is considerable overlap with efforts to monitor system strengthening itself.

^{xviii} www.washbat.org

^{xix} The tool facilitates the scoring of 'criteria', which together provide a score for an overarching 'governance function'. In most cases, several governance functions are grouped under each SWA building block. 'A criterion is like an indicator, stated in a positive way, that when progress is made on it, it contributes to achieving sector outcomes.' (UNICEF, 2018, p29).

2017-2018	<p>Various agencies continue to develop their own building block frameworks and use these for assessments in various countries:</p> <ul style="list-style-type: none"> • In 2017 and 2018, Water for People apply a sustainable services checklist in 26 district-level WASH programmes across nine countries.⁸³ This is one of the few examples where more than one baseline assessment has been conducted and there is now ongoing monitoring data available. Several changes were identified in the eight component indicators (effectively the building blocks) and 34 underlying metrics (individual, measurable indicators). • Water and Sanitation for the Urban Poor's framework, released in 2018, comprises 21 indicators, grouped into seven categories and varying for water vs. sanitation, to represent the organisation's 'understanding of a functional sector and the building blocks required'.⁸⁴ The resulting framework has been applied to undertake 'baseline sector functionality assessments' in six countries. • IRC has identified nine building blocks of the WASH system, developing and applying its thinking over the last decade.⁸⁵ It has undertaken a baseline in its six programme countries at national and municipal/district levels using a question and scoring system to assess the 'strength' of each building block to be scored.⁸⁶ • WaterAid has developed a framework of seven building blocks, which informs its process for designing programmes focused on sector strengthening, principally at district or city level. Among various participatory exercises, participants are encouraged to discuss and identify how the state of the building blocks in the district or city compare to a set of statements, which describe the 'typical' state of the building block in four contexts ('emergency', 'fragile but strengthening', 'transitional' and 'fully transitioned').⁸⁷ • The World Bank has continued to adapt building block frameworks, for example, using five building blocks to assess and score the enabling environment for sustainability of different models for rural water and sanitation services, in 16 countries globally⁸⁸ and then in seven countries in the Danube region.⁸⁹
-----------	--

Annex 3: Principle themes in international WASH building block frameworks

	Policy & Strategy	Institutional clarity	Accountability	Financing	Planning	Equity	Monitoring, evaluation & learning	Service delivery	Markets	Capacity	Political leadership	Decentralization	Social norms
CSO/ SDA/ MAPAS	Institutional framework Policy			Institutional framework Budget Sector development Expenditure	Institutional framework Planning	Sector development Equity		Sector development Output Sustainability Maintenance (water supply only) Sustainability Expansion (water supply only) Sustainability Up-take (sanitation only) Sustainability Use (sanitation only)	Sustainability Markets (sanitation only)				
WASHBat	Sector policy & strategy Sector policy & strategy	Institutional arrangements Coordination Institutional arrangements Service delivery arrangements	Institutional arrangements Accountability & regulation	Budgeting & financing Budget & expenditure Budgeting & financing Financing	Planning, monitoring & review Planning		Planning, monitoring & review Monitoring, evaluation & learning	Service providers Service providers		Capacity development Capacity development	Broader enabling environment Political leadership	Broader enabling environment Decentralization	Broader enabling environment Social norms
GLAAS	Governance Human rights to water and sanitation Governance Policy and plan development processes and effectiveness Governance National policies and plans Governance National target-setting process Governance National targets	Governance Coordination between actors Governance Coordinating with development partners Governance Institutional roles and responsibilities and lead agencies	Governance National regulations and standards Monitoring Regulation of drinking-water and sanitation/wastewater services Monitoring Functions of drinking-water regulators Monitoring Functions of sanitation/wastewater regulators Governance Community and user participation	Finance Budgets for WASH Finance Financial reporting Finance Absorption of external funds Finance Domestic absorption Finance Financing plan Finance Cost recovery strategies Finance External funding Finance Sufficiency of financing Finance Financial flows	Governance National development plans	Finance Equity Finance Affordability Governance Vulnerable groups in national WASH policies and plans Governance Vulnerable groups in national WASH targets	Monitoring National assessments and joint sector reviews (JSRs) Monitoring Use of monitoring data Monitoring Management information systems Monitoring Monitoring national targets Monitoring Monitoring vulnerable populations Monitoring Performance indicators Monitoring Drinking-water quality surveillance Monitoring Wastewater effluent surveillance	Governance Progress on national targets		Human resources Human resources needs assessments Human resources WASH training institutions/programmes Human resources Human resources for WASH operations and development			

www.washmatters.wateraid.org

WaterAid is a registered charity: Australia: ABN 99 700 687 141. Canada: 119288934 RR0001. India: U85100DL2010NPL200169. Japan: 特定非営利活動法人ウォーターエイドジャパン(認定NPO法人) WaterAid Japan is a specified non-profit corporation (certified NPO corporation). Sweden: Org.nr: 802426-1268, PG: 90 01 62-9, BG: 900-1629. UK: 288701 (England and Wales) and SC039479 (Scotland). US: WaterAid America is a 501(c) (3) non-profit organization

References

- ¹ United Nations (no date). *Sustainable Development Goal 6*. Available at: <https://sustainabledevelopment.un.org/sdg6> (accessed 24 January 2019)
- ² Huston A and Moriarty P (2018). *Building strong WASH systems for the SDGs: Understanding the WASH system and its building blocks*, p 18. IRC, The Netherlands.
- ³ WaterAid (2016). *Monitoring results of complex systems change*. WaterAid, UK.
- ⁴ Sanitation and Water for All (SWA) (2016). *Five building blocks*. SWA, US.
- ⁵ Velleman Y, Slaymaker T and Northover H (2009). 'Aid Compacts' built around national plans: How the Global Framework for Action on Sanitation can support country-led processes. WaterAid, UK.
- ⁶ de la Harpe J (2010). *Sector wide approach in the water sector*. IRC, The Netherlands.
- ⁷ Water for People (2017). *Monitoring framework*. Water for People, US.
- ⁸ Huston A and Moriarty P (2018). *Building strong WASH systems for the SDGs: Understanding the WASH system and its building blocks*. IRC, The Netherlands.
- ⁹ WaterAid (2018). *Sector strengthening programme design toolkit*. WaterAid, UK.
- ¹⁰ Norman G (2018). *An evaluative framework for urban WASH sector functionality*. WSUP, UK.
- ¹¹ See, for example, Federal Ministry for Economic Cooperation and Development (BMZ) (2017). *BMZ Water strategy*. BMZ, Germany; and DGIS (2016). *WASH strategy 2016-2030: Contributing to water, sanitation and hygiene for all, forever*. Ministry of Foreign Affairs, The Netherlands.
- ¹² WaterAid (2018). *Sector strengthening programme design toolkit*. WaterAid, UK.
- ¹³ Huston A and Moriarty P (2018). *Building strong WASH systems for the SDGs: Understanding the WASH system and its building blocks*. IRC, The Netherlands.
- ¹⁴ World Health Organization (WHO) (2010). *Monitoring the building blocks of health systems: A handbook of indicators and their measurement strategies*. WHO, Geneva.
- ¹⁵ Measure Evaluation (2010). *Health system strengthening*. Available at: www.measureevaluation.org/prh/rh_indicators/health-systems/hss (accessed 21 December 2018)
- ¹⁶ Adam T and Savigny D (2012). *Systems thinking for strengthening health systems in LMICs: Need for a paradigm shift*. Health Policy and Planning, vol 27, no 4, pp iv1-iv3.
- ¹⁷ Samuels F, Amaya A and Balabanova D (2017). *Drivers of health system strengthening: Learning from implementation of maternal and child health programmes in Mozambique, Nepal and Rwanda*. Health Policy and Planning, vol 32, no 7, pp 1015-1031.
- ¹⁸ Ibid.
- ¹⁹ Kruk M, Gage A, Arsenault C, et al (2018). *High-quality health systems in the Sustainable Development Goals era: Time for a revolution*. The Lancet Global Health Commission, vol 6, no 11, pp 1196-1252.
- ²⁰ Health Systems Governance Collaborative (2018). *Background paper for actionable governance – missing link meeting*. US.
- ²¹ Andrews M, Woolcock M and Pritchett L (2017). *Building state capability: Evidence, analysis, action*. Oxford University Press, UK.
- ²² Ramalingam B and Jones H (2008). *Exploring the science of complexity: Ideas and implications for development and humanitarian efforts*, p 8. ODI, UK.
- ²³ Mason N and Terefe B (2016). *Beyond political commitment to sanitation: Navigating incentives for prioritisation and course correction in Ethiopia*. WaterAid, UK.
- ²⁴ Ramalingam B and Jones H (2008). *Exploring the science of complexity: Ideas and implications for development and humanitarian efforts*, p 8. ODI, UK.
- ²⁵ AMCOW (African Ministers Council on Water) (2011). *AMCOW Country Status Overviews – Regional synthesis report. Pathways to Progress: Transitioning to country-led service delivery pathways to meet Africa's water supply and sanitation targets*. The World Bank, US.
- ²⁶ DGIS (2016). *WASH strategy 2016-2030: Contributing to water, sanitation and hygiene for all, forever*, p 7. Ministry of Foreign Affairs, The Netherlands.
- ²⁷ Jimenez A, Jawara D, LeDeunff H, et al (2017). *Sustainability in practice: Experiences from rural water and sanitation services in West Africa*. Sustainability, vol 9, no 3, p 403.
- ²⁸ Andrews M, Woolcock M and Pritchett L (2017). *Building state capability: Evidence, analysis, action*, p 5. Oxford University Press, UK.

- 29 Wild L, Booth D, Cummings C, et al (2015). *Adapting development: Improving services to the poor*. ODI Report, UK.
- 30 Wild L and Ramalingam B (2018). *Building a global learning alliance on adaptive management*. ODI Report, UK.
- 31 Eberhard R (2018). *Access to water and sanitation in sub-Saharan Africa*. GIZ, Germany.
- 32 Huston A and Moriarty P (2018). *Building strong WASH systems for the SDGs: Understanding the WASH system and its building blocks*. IRC, The Netherlands.
- 33 UNICEF (2018). *Water, sanitation and hygiene bottleneck analysis tool (WASH BAT) user guide*. Available at www.washbat.org/wp-content/uploads/en/WASH_BAT_2_User_Manual.pdf (accessed 24 January 2019).
- 34 Ramalingam B and Jones H (2008). *Exploring the science of complexity: Ideas and implications for development and humanitarian efforts*, p 8. ODI, UK.
- 35 WaterAid (2016). *Monitoring results of complex systems change*. WaterAid, UK.
- 36 Global Obesity Prevention Centre (2016). *Spaces MERL: Systems and complexity white paper*. USAID, US.
- 37 Sustainable WASH Systems Learning Partnership (2018). *Network analysis and systems assessment in the rural sanitation and hygiene sector in Cambodia*. USAID, US.
- 38 McNicholl D (2018). *Kabarole WASH sustainability actor and factor analysis*. IRC and Aguaconsult.
- 39 USAID (2016). *SPACES-MERL: Systems and Complexity White Paper*, p10. USAID, US.
- 40 WaterAid (2018). *Sector strengthening programme design toolkit*. WaterAid, UK.
- 41 WaterAid (2016). *Monitoring results of complex systems change*. WaterAid, UK.
- 42 DFID (2018). *Urban sanitation policy programme*. [online] Available at <https://devtracker.dfid.gov.uk/projects/GB-1-204701/documents> (accessed 23 January 2019).
- 43 ODI (2009). *Tools for knowledge and learning: A guide for development and humanitarian organisations: Most significant change (MSC)*. ODI, UK.
- 44 Willetts J and Crawford P (2007). *The most significant lessons about the most significant change technique*. Development in Practice, vol 17, no 3, pp 367-379.
- 45 UNICEF (2018). *Strengthening enabling environment for water, sanitation and hygiene (WASH) guidance note*. UNICEF, US.
- 46 World Bank (2017a). *Reducing inequalities in water supply, sanitation and hygiene in the era of the sustainable development goals: Synthesis report of the WASH Poverty Diagnostic Initiative*. World Bank Group, US.
- 47 AMCOW (African Ministers Council on Water) (2011). *AMCOW Country Status Overviews – Regional synthesis report. Pathways to progress: Transitioning to country-led service delivery pathways to meet Africa's water supply and sanitation targets*, p7. The World Bank, US.
- 48 GIZ (2018a). *Access to water and sanitation in sub-Saharan Africa: Review of sector reforms and investments, and key findings to inform future support to sector development*. GIZ, Germany.
- 49 Marks S, Kumpel E, Guo J, et al (2018). *Pathways to sustainability: A fuzzy-set qualitative comparative analysis of rural water supply programmes*. Journal of Cleaner Production, vol 205, no 12, pp 789-798.
- 50 Ibid; Clist P (2017). *What works for payment by results mechanisms in DFID programmes*. DFID, UK.
- 51 Birdsall N, Mahgoub A and Savedoff W (2010). *Cash on delivery: A new approach to foreign aid*. Center for Global Development, US.
- 52 WSUP (2011). *Progress-linked finance: A study of the feasibility and practicality of a proposed WASH financing approach*. WSUP, UK; GIZ (2018b). *Finance concept note*, unpublished.
- 53 Ibid.
- 54 World Bank (2017b). *Improving service levels and impact on the poor: A diagnostic of water supply, sanitation, hygiene and poverty in Indonesia*. World Bank, US.
- 55 Willetts J, Howard M, Edwards J, et al (2017). *Water and sanitation hibah: Independent review*. DFAT, Australia.
- 56 Ibid.
- 57 Setiono I (2017). *Indonesia – National Urban Water Supply Project (NUWSP): Environmental assessment: Environmental and social management framework*. World Bank, Indonesia.
- 58 Chong Y, Abeyesuriya K, Hidayat L, et al (2016). *Strengthening local governance arrangements for sanitation: Case studies of small cities in Indonesia*. Aquatic Procedia, vol 6, no 8, pp 64-73.

-
- ⁵⁹ Government of Ethiopia (2018a). *One WASH National Program: A multi-sectoral SWAP*. Phase 2 Program Document. GOE, Ethiopia.
- ⁶⁰ Mason N and Terefe B (2016b). *Beyond political commitment to sanitation: Navigating incentives for prioritisation and course correction in Ethiopia*. WaterAid, UK.
- ⁶¹ Government of Ethiopia (2018b). *One WASH National Program: A multi-sectoral SWAP*. Review of Phase 1. GOE, Ethiopia.
- ⁶² Government of Ethiopia (2016). *One WASH National Program Annual Report 2008 EFY*. GOE, Ethiopia; Government of Ethiopia (2018). *One WASH National Program (OWNP) Annual Report 2009 EFY*. GOE, Ethiopia.
- ⁶³ Government of Ethiopia (2018a). *One WASH National Program: A multi-sectoral SWAP*. Phase 2 Program Document. GOE, Ethiopia.
- ⁶⁴ Mason N and Terefe B (2016). *Beyond political commitment to sanitation: Navigating incentives for prioritisation and course correction in Ethiopia*. WaterAid, UK.
- ⁶⁵ Government of Ethiopia (2018a). *One WASH National Program: A multi-sectoral SWAP*. Phase 2 Program Document. GOE, Ethiopia.
- ⁶⁶ IRC WASH and WaterAid Ethiopia (2018). *WASH initiatives taking a Woreda-wide systems strengthening approach to improving service delivery*. Background note, unpublished.
- ⁶⁷ Adank et al (2018). *SWS Ethiopia baseline study*. Forthcoming. Sustainable WASH Systems Learning Partnership, Ethiopia.
- ⁶⁸ Government of Ethiopia (2018b). *One WASH National Program: A multi-sectoral SWAP*. Review of Phase 1. GOE, Ethiopia; Government of Ethiopia (2015). *Health Sector Transformation Plan*. Ministry of Health, Ethiopia.
- ⁶⁹ Government of Ethiopia (2017). *Health Sector Transformation Plan-1 Annual performance report*. Ministry of Health, Ethiopia.
- ⁷⁰ Government of Ethiopia (2015). *Ethiopia Health Sector Transformation Plan*. Ministry of Health, Ethiopia.
- ⁷¹ Ibid.
- ⁷² Lavers T (2016). *Social protection in an aspiring 'developmental state': The political drivers of community-based health insurance in Ethiopia*. The University of Manchester, UK.
- ⁷³ AMCOW (African Ministers Council on Water) (2011). *AMCOW Country Status Overviews – Regional synthesis report. Pathways to progress: Transitioning to country-led service delivery pathways to meet Africa's water supply and sanitation targets*. The World Bank, US.
- ⁷⁴ UN Water and World Health Organization (2008). *UN-Water global annual assessment of sanitation and drinking water*. WHO, Geneva.
- ⁷⁵ Ibid.
- ⁷⁶ Velleman Y, Slaymaker T and Northover H (2009). *'Aid Compacts' built around national plans: How the Global Framework for Action on Sanitation can support country-led processes*. WaterAid, UK.
- ⁷⁷ AMCOW (African Ministers Council on Water) (2011). *AMCOW Country Status Overviews – Regional synthesis report. Pathways to progress: Transitioning to country-led service delivery pathways to meet Africa's water supply and sanitation targets*, p7. The World Bank, US.
- ⁷⁸ World Bank (2015a). *Turning finance into services for the future: A regional synthesis of the service delivery assessments for water supply and sanitation*. World Bank, US.
- ⁷⁹ World Bank (2015b). *MAPAS synthesis note*. World Bank, US.
- ⁸⁰ Sanitation and Water for All (SWA) (2016). *Five building blocks*. Available at: sanitationandwaterforall.org/wp-content/uploads/download-manager-files/building%20blocks%20%20pager.pdf (accessed 24 January 2019).
- ⁸¹ Jiménez A, Le Deunff H, Avello P, et al (2016). *Enabling environment: Accountability for sustainability*. UNICEF, SIWI Water Governance Facility & UNDP.
- ⁸² UNICEF (no date). *WASH Bottleneck Analysis Tool: Country implementation guide*. UNICEF, US.
- ⁸³ Water for People (2017). *Monitoring framework*. Water for People, US.
- ⁸⁴ Norman G (2018). *An evaluative framework for urban WASH sector functionality*, p1. WSUP, UK.
- ⁸⁵ Huston A and Moriarty P (2018). *Building strong WASH systems for the SDGs: Understanding the WASH system and its building blocks*. IRC, The Netherlands.
- ⁸⁶ Smits S (2018). *Weekly WASH graph: A sneak preview at measuring system strength*. Available at ircwash.org/blog/weekly-wash-graph-sneak-preview-measuring-system-strength (accessed 24 January 2019).
- ⁸⁷ WaterAid (2018). *Sector strengthening programme design toolkit*. WaterAid, UK.
-

⁸⁸ World Bank (2018a). *Sustainability assessment of rural water service delivery models: Findings of a multi-country review*. World Bank, US.

⁸⁹ World Bank (2018b). *A review of rural water and sanitation services in seven countries of the Danube Region*. World Bank, US.